

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
29. Juli 2004 (29.07.2004)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2004/063358 A1

(51) Internationale Patentklassifikation: C12N 1/15,
15/80, C12P 23/00

(21) Internationales Aktenzeichen: PCT/EP2004/000100

(22) Internationales Anmeldedatum:
9. Januar 2004 (09.01.2004)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
103 00 649.4 9. Januar 2003 (09.01.2003) DE
103 41 272.7 8. September 2003 (08.09.2003) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme
von US): BASF AKTIENGESELLSCHAFT [DE/DE];
67056 Ludwigshafen (DE).

(72) Erfinder; und
(75) Erfinder/Anmelder (nur für US): MATUSCHEK,
Markus [DE/DE]; Karolinenstr. 5, 69469 Weinheim
(DE). HEINEKAMP, Thorsten [DE/DE]; Alte Ziegelei
1h, 38419 Hannover (DE). SCHMIDT, Andre [DE/DE];
Magdeburger Str. 11, 31832 Springe (DE). BRAKHAGE,
Axel [DE/DE]; Schneiderberg 58, 38167 Hannover (DE).

(74) Anwalt: FITZNER, Uwe; Lintorfer Strasse 10, 40878
Ratingen (DE).

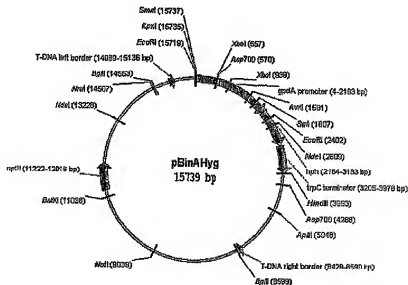
(81) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verfügbare nationale Schutzrechtsart): AE, AG, AL,
AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH,
CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES,
FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,
MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD FOR THE GENETIC MODIFICATION OF ORGANISMS OF THE GENUS BLAKESLEA, CORRE-
SPONDING ORGANISMS, AND THE USE OF THE SAME

(54) Bezeichnung: VERFAHREN ZUR GENTECHNISCHEN VERÄNDERUNG VON ORGANISMEN DER GATTUNG BLA-
KESLEA, ENTSPRECHENDE ORGANISMEN UND DEREN VERWENDUNG

VEKTOR pBinAHyg



(57) Abstract: The invention relates to a method for producing a genetically modified organism of the genus *Blakeslea*, said method comprising the following steps: (i) at least one of the cells is transformed, (ii) the cells obtained in step (i) are optionally rendered homokaryotic, so that cells are created in which the nuclei are all homogeneously modified in at least one genetic characteristic and convert said genetic modification into an expression, and (iii) the genetically modified cell or cells are selected and cultivated.

[Fortsetzung auf der nächsten Seite]

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
29. Juli 2004 (29.07.2004)

PCT

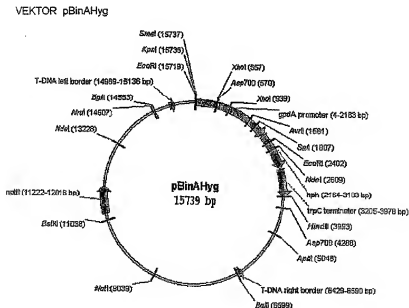
(10) Internationale Veröffentlichungsnummer
WO 2004/063358 A1

- | | | | |
|---|---|--|---|
| (51) Internationale Patentklassifikation*:
15/80, C12P 23/00 | C12N 1/15, | (72) Erfinder: und | (75) Erfinder/Anmelder (nur für US): |
| (21) Internationales Aktenzeichen: | PCT/EP2004/000100 | MATUSCHEK,
Markus [DE/DE]; Karolinenstr. 5, 69469 Weinheim
(DE), HEIDENKAMP, Thorsten [DE/DE]; Alte Ziegelei
1b, 38419 Hannover (DE), SCHMIDT, Andre [DE/DE];
Magdeburger Str. 11, 31832 Springe (DE), BRAKHAGE,
Axel [DE/DE]; Schneiderberg 58, 38167 Hannover (DE). | |
| (22) Internationales Anmeldedatum: | 9. Januar 2004 (09.01.2004) | | |
| (25) Einreichungssprache: | Deutsch | (74) Anwalt: | FITZNER, Uwe; Lintorfer Strasse 10, 40878
Ratingen (DE). |
| (26) Veröffentlichungssprache: | Deutsch | | |
| (30) Angaben zur Priorität: | 103 09 649 4 9. Januar 2003 (09.01.2003) DE
103 41 272 7 8. September 2003 (08.09.2003) DE | | (81) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verifizierbare nationale Schutzrechtsart): AE, AG, AL,
AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH,
CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES,
FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MG,
MK, MN, MW, MX, MY, MZ, NA, NI, NO, NZ, OM, PG, |
| (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme
von US): | BASF AKTIENGESellschaft [DE/DE];
67056 Ludwigshafen (DE). | | |

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD FOR THE GENETIC MODIFICATION OF ORGANISMS OF THE GENUS BLAKESLEA, CORRESPONDING ORGANISMS, AND THE USE OF THE SAME

(54) **Bezeichnung:** VERFAHREN ZUR GENTECHNISCHEN VERÄNDERUNG VON ORGANISMEN DER GATTUNG BLAKESLEA. ENTSPRECHENDE ORGANISMEN UND DEREN VERWENDUNG



- (57) **Abstract:** The invention relates to a method for producing a genetically modified organism of the genus *Blakeslea*, said method comprising the following steps: (i) at least one of the cells is transformed, (ii) the cells obtained in step (i) are optionally rendered homokaryotic, so that cells are created in which the nuclei are all homogeneously modified in at least one genetic characteristic and convert said genetic modification into an expression, and (iii) the genetically modified cell or cells are selected and cultivated.

[Fortsetzung auf der nächsten Seite]

WO 2004/063358 A1



PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

— vor Ablauf der für Änderungen der Ansprüche geltenden Frist; Veröffentlichung wird wiederholt, falls Änderungen eintreffen

- (84) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

Via öffentlich:

— mit internationalem Recherchenbericht

(57) Zusammenfassung: Verfahren zur gentechnischen Veränderung von Organismen der Gattung *Blakeslea*, entsprechende Organismen und deren Verwendung Zusammenfassung Verfahren zur Herstellung eines gentechnisch veränderten Organismus der Gattung *Blakeslea* umfassend (i) Transformation mindestens einer der Zellen, (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder mehreren genetischen Merkmalen alle gleichzeitig verändert sind und diese gentechnische Veränderung zur Anspröng bringen, und (iii) Selektion und Anzucht der gentechnisch veränderten Zelle oder Zellen.

Verfahren zur gentechnischen Veränderung von Organismen der
Gattung Blakeslea, entsprechende Organismen und deren
Verwendung

Die Erfindung betrifft ein Verfahren zur gentechnischen Veränderung von Organismen der Gattung Blakeslea, entsprechende Organismen und deren Verwendung.

Pilze der Gattung Blakeslea sind als Produktionsorganismen bekannt. So wird z. B. Blakeslea trispora als Produktionsorganismus für β -Carotin (Ciegler, 1965, Adv Appl Microbiol. 7:1) und Lycopin verwendet (EP 1201762, EP 1184464, WO 03/038064). Daneben kommt Blakeslea zur Produktion anderer lipophiler Substanzen in Frage wie z.B. andere Carotinoide und deren Vorstufen, Phospholipide, Triacylglyceride, Steroide, Wachse, fettlösliche Vitamine, Provitamine und Cofaktoren oder zur Produktion hydrophiler Substanzen wie z.B. Eiweiße, Aminosäuren, Nukleotide und wasserlösliche Vitamine, Provitamine und Cofaktoren.

Die hohen Produktivitäten für β -Carotin und Lycopin machen Blakeslea, insbesondere Blakeslea trispora attraktiv für die wirtschaftliche fermentative Herstellung von Carotinoiden und deren Vorstufen.

Allerdings ist es auch von Interesse, die Produktivitäten der bisher natürlicherweise produzierten Carotine und deren Vorstufen weiter zu steigern und die Herstellung weiterer Carotinoide, wie z. B. Xanthophylle zu ermöglichen, die von Blakeslea bisher nicht oder nur in sehr geringem Maße gebildet und isoliert werden können.

Carotinoide werden Futtermitteln, Nahrungsmitteln, Nahrungsergänzungsmitteln, Kosmetika und Arzneimitteln zugesetzt. Die Carotinoide dienen vor allem als Pigmente zur Färbung. Daneben werden

die antioxidative Wirkung der Carotinoide und andere Eigenschaften dieser Substanzen genutzt. Man unterteilt die Carotinoide in die reinen Kohlenwasserstoffe, die Carotine und die sauerstoffhaltigen Kohlenwasserstoffe, die Xanthophylle. Xanthophylle wie Canthaxanthin
5 und Astaxanthin werden beispielsweise zur Pigmentierung von Hühnerelern und Fischen eingesetzt (Britton et al. 1998, Carotinoids, Vol 3, Biosynthesis and Metabolism). Die Carotine β -Carotin und Lycopin werden vor allem in der Humanernährung eingesetzt. β -Carotin wird beispielsweise als Getrnkefarbstoff verwendet. Lycopin hat eine
10 krankheitsvorbeugende Wirkung (Argwal und Rao, 2000, CMAJ 163:739-744; Rao und Argwal 1999, Nutrition Research 19:305-323). Die farblose Carotinoidvorstufe Phytoen kommt vor allem fr Anwendungen als Antioxidans in Frage.

15 Der berwiegende Teil der Carotinoide und deren Vorstufen, die als Zusatzstoffe fr die oben genannten Anwendungen eingesetzt werden, wird durch chemische Synthese hergestellt. Die chemische Synthese ist mehrstufig, technisch sehr aufwendig und verursacht hohe Herstellkosten. Fermentative Verfahren sind demgegenber technisch verhltnismig
20 einfach und basieren auf kostengnstigen Einsatzstoffen. Fermentative Verfahren zur Herstellung von Carotinoiden knnen dann wirtschaftlich attraktiv und wettbewerbsfhig zur chemischen Synthese sein, wenn die Produktivitt der bisherigen fermentativen Verfahren gesteigert wrde oder neue Carotinoide auf Basis der bekannten Produktionsorganismen
25 hergestellt werden knnten.

Ein Verfahren zur gentechnischen Vernderung von *Blakeslea trispora* ist erforderlich insbesondere, wenn *Blakeslea* zur Herstellung von Xanthophyllen genutzt werden soll, weil diese Verbindungen
30 natrlicherweise von *Blakeslea* nicht synthetisiert werden.

Von *Blakeslea trispora* sind bisher verschiedene DNA-Sequenzen bekannt. Insbesondere die DNA-Sequenz, die für die Gene der Carotinoidbiosynthese von Geranylgeranylpyrophosphat bis β -Carotin codiert (WO 03/027293).

5

Allerdings sind bisher keine Methoden zur gentechnischen Veränderung von *Blakeslea*, insbesondere *Blakeslea trispora* bekannt.

Als Methode zur Herstellung von gentechnischen veränderten Pilzen wurde in einigen Fällen die *Agrobacterium*-vermittelte Transformation erfolgreich eingesetzt. So sind z. B. folgende Organismen durch *Agrobakterien* transformiert worden: *Saccharomyces cerevisiae* (Bundock et al., 1995, EMBO Journal, 14:3206–3214), *Aspergillus awamori*, *Aspergillus nidulans*, *Aspergillus niger*, *Colletotrichum gloeosporioides*,
15 *Fusarium solani* pisi, *Neurospora crassa*, *Trichoderma reesei*, *Pleurotus ostreatus*, *Fusarium graminearum* (van der Toorren et al., 1997, EP 870835), *Agraricus bisporus*, *Fusarium venenatum* (de Groot et al., 1998, Nature Biotechnol. 16:839–842), *Mycosphaerella graminicola* (Zwiars et al. 2001, Curr. Genet. 39:388–393), *Glarea lozoyensis* (Zhang et al., 2003, Mol. Gen. Genomics 268:645–655), *Mucor miehei* (Monfort et al. 2003, FEMS Microbiology Lett. 244:101 – 106).
20

Von Interesse ist besonders eine homologe Rekombination, bei der zwischen der einzuführenden DNA und der Zell-DNA möglichst viele Sequenzhomologien bestehen, so dass eine ortsspezifische Einführung bzw. Ausschaltung von genetischer Information im Genom des Empfängerorganismus möglich ist. Andernfalls wird die Spender-DNA durch illegitime bzw. nicht-homologe Rekombination ins Genom des Empfängerorganismus integriert, was nicht ortsspezifisch erfolgt.

30

Eine durch *Agrobacterium* vermittelte Transformation und anschließende homologe Rekombination der transferierten DNA wurde bisher bei folgenden Organismen nachgewiesen: *Aspergillus awamori* (Gouka et al. 1999, *Nature Biotech* 17:598-601), *Glarea lozoyensis* (Zhang et al., 2003, 5 *Mol. Gen. Genomics* 268:645-655), *Mycosphaerella graminicola* (Zwiers et al. 2001, *Curr. Genet.* 39:388-393).

Als weitere Methode zur Transformation von Pilzen ist die Elektroporation bekannt. Die integrative Transformation von Hefe durch Elektroporation 10 wurde von Hill, *Nucl. Acids. Res.* 17:8011 gezeigt. Für filamentöse Pilze wurde die Transformation durch Chakaborty und Kapoor beschrieben (1990, *Nucl. Acids. Res.* 18:6737).

Eine „biolistische“ Methode, d.h. die Übertragung von DNA durch 15 Beschuss von Zellen mit DNA-beladenen Partikeln wurde beispielsweise für *Trichoderma harzianum* und *Gliocladium virens* beschrieben (Lorito et al. 1993, *Curr. Genet.* 24:349-356).

Diese Methoden konnten bisher jedoch nicht erfolgreich zur gezielten 20 genetischen Veränderung von *Blakeslea* und insbesondere *Blakeslea trispora* eingesetzt werden.

Eine besondere Schwierigkeit bei der Herstellung von gezielt genetisch veränderten *Blakeslea* und *Blakeslea trispora* ist die Tatsache, dass deren 25 Zellen in allen Stadien des sexuellen und des vegetativen Zellzyklus mehrkernig sind. In Sporen von *Blakeslea trispora* Stamm NRRL2456 und NRRL2457 wurden z. B. im Durchschnitt 4,5 Kerne pro Spore nachgewiesen (Metha und Cerdá-Olmedo, 1995, *Appl. Microbiol. Biotechnol.* 42:836-838). Dies hat zur Folge, dass die gentechnische 30 Veränderung in aller Regel nur in einem oder wenigen Kernen vorliegt, die Zellen also heterokaryotisch sind.

- Sollen die genetisch veränderten *Blakeslea*-Arten, insbesondere *Blakeslea trispora* zur Produktion eingesetzt werden, so ist es insbesondere bei einer Gendelektion wichtig, dass in den Produktionsstämmen die gentechnische Veränderung in allen Kernen vorliegt, so dass eine stabile und hohe Syntheseleistung ohne Nebenprodukte möglich wird. Die Stämme müssen folglich in Bezug auf die gentechnische Veränderung homokaryotisch sein.
- 10 Lediglich für *Phycomyces blakesleeanus* ist ein Verfahren beschrieben worden, um homokaryotische Zellen zu erzeugen (Roncero et al., 1984, Mutat. Res. 125:195). Durch Zugabe des mutagenen Agens MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) werden nach dem dort beschriebenen Verfahren Kerne in den Zellen eliminiert, so dass statistisch eine gewisse
- 15 Anzahl von Zellen mit nur noch einem funktionellem Kern vorliegt. Die Zellen werden dann einer Selektion unterzogen, in der nur einkernige Zellen mit einem rezessiven Selektionsmarker zu einem Mycel auswachsen können. Die Nachkommen dieser selektierten Zellen sind mehrkernig und homokaryotisch. Ein rezessiver Selektionsmarker für
- 20 *Phycomyces blakesleeanus* ist z. B. dar^+ -Stämme nehmen das toxische Riboflavin-Analog 5-Carbon-5-deazariboflavin auf; dar^- -Stämme dagegen nicht (Delbrück et al. 1979, Genetics 92:27). Die Selektion von rezessiven Mutanten erfolgt durch Zugabe von 5-Carbon-5-deazariboflavin (DARF).
- 25 Allerdings ist dieses Verfahren nicht für *Blakeslea*, insbesondere *Blakeslea trispora* bekannt und insbesondere nicht im Zusammenhang mit einer Transformation beschrieben worden.
- 30 Aufgabe der vorliegenden Erfindung ist es, ein Verfahren bereitzustellen, mit dem eine gentechnische Veränderung von *Blakeslea*-Stämmen,

- insbesondere *Blakeslea trispora* möglich ist. Darüber hinaus ist es Aufgabe der Erfindung ein Verfahren bereitzustellen, das die Herstellung homokaryotischer genetisch veränderter Stämme erlaubt. Ferner ist es eine Aufgabe der Erfindung entsprechend gentechnisch veränderte Zellen bereitzustellen.

Diese Aufgabe wird durch ein Verfahren zur Herstellung eines gentechnisch veränderten Organismus der Gattung *Blakeslea* gelöst, umfassend

- 10 (i) Transformation mindestens einer der Zellen,
- (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder in mehreren genetischen Merkmalen alle gleichartig verändert sind und diese genetische Veränderung zur Ausprägung bringen, und
- 15 (iii) Selektion der gentechnisch veränderten Zelle oder Zellen.

Mit der erfindungsgemäßen Methode ist es möglich, mehrkernige Zellen der Pilze *Blakeslea* gezielt und stabil genetisch zu verändern, um so Mycel aus Zellen mit einheitlichen Kernen zu gewinnen. Vorzugsweise handelt es sich um Zellen von Pilzen der Art *Blakeslea trispora*.

Unter Transformation wird die Übertragung einer genetischen Information in den Organismus, insbesondere Pilz, verstanden. Darunter sollen alle dem Fachmann bekannten Möglichkeiten zur Einschleusung der Information, insbesondere DNA, fallen, z. B. Beschuss mit DNA-beladenen Partikeln, Transformation mittels Protoplasten, Mikroinjektion von DNA, Elektroporation, Konjugation oder Transformation kompetenter Zellen, Chemikalien oder Agrobakterien vermittelte Transformation. Als genetische Information werden ein Genabschnitt, ein Gen oder mehrere

Gene verstanden. Die genetische Information kann z. B. mit Hilfe eines Vectors oder als freie Nukleinsäure (z. B. DNA, RNA) und auf sonstige Weise in die Zellen eingebracht und entweder durch Rekombination ins Wirtsgenom eingebaut oder in freier Form in der Zelle vorliegen.

5 Besonders bevorzugt ist hierbei die homologe Rekombination.

Bevorzugte Transformationsmethode ist die *Agrobacterium tumefaciens*-vermittelte Transformation. Hierzu wird zunächst die zu transferierende Spender-DNA in einen Vektor eingefügt, der (i) flankierend zu der zu transferierenden DNA die T-DNA-Enden trägt, der (ii) einen Selektionsmarker enthält und der (iii) ggf. Promotoren und Terminatoren für die Genexpression der Spender-DNA aufweist. Dieser Vektor wird in einen *Agrobacterium tumefaciens*-Stamm übertragen, der ein Ti-Plasmid mit den vir-Genen enthält. vir-Gene sind für den DNA-Transfer in *Blakeslea* verantwortlich. Mit diesem Zwei-Vektor-System wird die DNA von *Agrobacterium* in *Blakeslea* übertragen. Hierzu werden die *Agrobakterien* zunächst in Gegenwart von Acetosyringone inkubiert. Acetosyringone induziert die vir-Gene. Anschließend werden Sporen von *Blakeslea trispora* zusammen mit den induzierten Zellen von *Agrobacterium tumefaciens* auf Acetosyringone-haltigem Medium inkubiert und dann auf Medium übertragen, das eine Selektion der Transformanten, d.h. der gentechnisch veränderten Stämme von *Blakeslea* ermöglicht.

Der Begriff Vector wird in der vorliegenden Anmeldung als eine Bezeichnung für ein DNA-Molekül verwendet, das zum Einschleusen und ggf. zur Vermehrung von Fremd-DNA in eine Zelle dient (siehe auch "Vector" in Römpp Lexikon Chemie – CDROM Version 2.0, Stuttgart/New York: Georg Thieme Verlag 1999). In der vorliegenden Anmeldung sollen unter dem Begriff "Vector" Plasmide, Cosmide usw. verstanden werden, die diesem Zweck dienen.

Unter Expression wird in der vorliegenden Anmeldung die Übertragung einer genetischen Information ausgehend von DNA oder RNA in ein Gen-Produkt (hier vorzugsweise Carotinoide) verstanden und soll auch den Begriff der Überexpression beinhalten, womit eine verstärkte Expression
5 gemeint ist, so dass ein bereits in der nicht transformierten Zelle (Wildtyp) hergestelltes Produkt verstärkt produziert wird oder einen großen Teil des gesamten Gehaltes der Zelle ausmacht.

Unter gentechnischer Veränderung soll die Einschleusung genetischer Information in einen Empfängerorganismus, so dass diese stabil exprimiert
10 und bei der Zellteilung weitergegeben wird, verstanden werden. Danach wird gegebenenfalls die Homokaryontisierung durchgeführt, d.h. die Herstellung von Zellen, die nur einheitliche Kerne enthalten, d. h. Kerne mit gleichem genetischem Informationsgehalt.

15 Diese Homokaryontisierung ist insbesondere notwendig, wenn die durch Transformation eingeführte genetische Information rezessiv vorliegt, d. h. nicht zur Ausprägung gelangt. Führt die Transformation aber zu einem dominanten Vorliegen der genetischen Information, d. h. wird sie
20 ausgeprägt, so ist eine Homokaryontisierung nicht unbedingt nötig.

Vorzugsweise wird zur Homokaryontisierung eine Selektion der einkernigen Sporen durchgeführt. Von Natur aus ist ein geringer Anteil der Sporen von *Blakeslea trispora* einkernig, so dass sich diese ggf. nach spezifischer
25 Markierung z. B. Färbung der Zellkerne aussortieren lassen. Dies wird bevorzugterweise mittels FACS (Fluorescence Activated Cell Sorting) anhand der geringeren Fluoreszenz der einkernigen Zellen durchgeführt.

Alternativ kann zur Homokaryontisierung zunächst eine Kernreduktion
30 durchgeführt werden. Hierzu kann ein mutagenes Agens eingesetzt werden, wobei es sich insbesondere um N-Methyl-N'-nitro-nitrosoguanidin

(MNNG) handelt. Auch die Verwendung von energiereichen Strahlen, wie UV- oder Röntgen-Strahlen zur Kernreduktion ist möglich. Anschließend kann zur Selektion auf das FACS Verfahren oder rezessive Selektionsmarker zurückgegriffen werden.

5

Unter Selektion wird die Auswahl von Zellen verstanden, deren Kerne dieselbe genetische Information beinhalten, d. h. Zellen die die gleichen Eigenschaften aufweisen, wie Resistenzen oder die Herstellung bzw. vermehrte Herstellung eines Produktes. In der Selektion werden neben
10 der FACS Methode bevorzugt 5-Carbon-5-deazariboflavin (darf) und Hygromycin (hyg) oder 5'-Fluororotat (FOA) und Uracil eingesetzt.

Der in der Transformation (i) eingesetzte Vector kann derart gestaltet sein, dass die im Vector enthaltene genetische Information in das Genom
15 mindestens einer Zelle integriert wird. Dabei kann genetische Information in der Zelle ausgeschaltet werden.

Der in der Transformation (i) eingesetzte Vector kann aber auch derart ausgestaltet sein, dass die im Vector enthaltene genetische Information in
20 der Zelle exprimiert wird, d. h. genetische Information eingefügt wird, die im korrespondierenden Wildtyp nicht vorhanden ist oder die durch die Transformation verstärkt bzw. überexprimiert wird.

Der Vector kann beliebige genetische Informationen zur genetischen
25 Veränderungen von Organismen der Gattung *Blakeslea* enthalten.

Unter „genetischer Information“ werden vorzugsweise Nukleinsäuren verstanden, deren Einbringung in den Organismus der Gattung *Blakeslea* zu einer genetischen Veränderung in Organismen der Gattung *Blakeslea*,
30 also beispielsweise zu einer Verursachung, Erhöhung oder Reduzierung von Enzymaktivitäten im Vergleich zum Ausgangsorganismus führen.

Der Vector kann beispielsweise genetische Information zur Herstellung lipophiler Substanzen enthalten wie z.B. Carotinoide und deren Vorstufen, Phospholipide, Triacylglyceride, Steroide, Wachse, fettlösliche Vitamine, Provitamine und Cofaktoren oder genetische Information zur Herstellung hydrophiler Substanzen wie z.B. Eiweiße, Aminosäuren, Nukleotide und wasserlösliche Vitaminen, Provitamine und Cofaktoren.

Bevorzugterweise enthält der eingesetzte Vector genetische Informationen zur Herstellung von Carotinoiden oder Xanthophyllen oder deren Vorstufen.

Bevorzugterweise enthält der Vektor genetische Information, die eine Lokalisierung der Carotinoidbiosynthese-Enzyme in dem Zellkompartiment bewirkt, in dem die Carotinoidbiosynthese stattfindet.

Besonders bevorzugt sind genetische Informationen zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3- und 3'-Hydroxyechinenon, Lycopin, Lutein, β -Carotin, Phytoen oder Phytoflu. Ganz besonders bevorzugt sind genetische Informationen zur Herstellung von Phytoen, Bixin, Lycopin, Zeaxanthin, Canthaxanthin und Astaxanthin.

Entsprechend werden in einer bevorzugten Variante der Erfindung Organismen hergestellt und kultiviert, die über eine erhöhte Syntheserate für Zwischenprodukte der Carotinoidbiosynthese verfügen und folglich eine erhöhte Produktivität für Endprodukte der Carotinoidbiosynthese aufweisen. Zur Erhöhung der Syntheserate für Zwischenprodukte der Carotinoidbiosynthese werden insbesondere die Aktivitäten der Enzyme 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A-Reduktase,

Isopentenylpyrophosphat-Isomerase und Geranylpyrophosphatsynthase gesteigert.

- 5 Entsprechend werden in einer besonders bevorzugten Variante der Erfindung Organismen hergestellt und kultiviert, die gegenüber dem Wildtyp eine erhöhte HMG-CoA-Reduktase-Aktivität aufweisen.

10 Unter HMG-CoA-Reduktase-Aktivität wird die Enzymaktivität einer HMG-CoA-Reduktase (3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A-Reduktase) verstanden.

Unter einer HMG-CoA-Reduktase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A in Mevalonat umzuwandeln.

- 15 Dementsprechend wird unter HMG-CoA-Reduktase-Aktivität die in einer bestimmten Zeit durch das Protein HMG-CoA-Reduktase umgesetzte Menge 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A bzw. gebildete Menge Mevalonat verstanden.

- 20 Bei einer erhöhten HMG-CoA-Reduktase-Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein HMG-CoA-Reduktase die umgesetzte Menge 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A bzw. die gebildete Menge Mevalonat erhöht.

- 25 Vorzugsweise beträgt diese Erhöhung der HMG-CoA-Reduktase-Aktivität mindestens 5%, weiter bevorzugt mindestens 20%, weiter bevorzugt mindestens 50%, weiter bevorzugt mindestens 100%, bevorzugter mindestens 300%, noch bevorzugter mindestens 500%, insbesondere mindestens 600% der HMG-CoA-Reduktase-Aktivität des Wildtyps.

In einer bevorzugten Ausführungsform erfolgt die Erhöhung der HMG-CoA-Reduktase-Aktivität gegenüber dem Wildtyp durch eine Erhöhung der Genexpression einer Nukleinsäure codierend eine HMG-CoA-Reduktase.

- 5 In einer besonders bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens erfolgt die Erhöhung der Genexpression einer Nukleinsäure codierend eine HMG-CoA-Reduktase indem man ein Nukleinsäurekonstrukt, enthaltend eine Nukleinsäure codierend eine HMG-CoA-Reduktase in den Organismus einbringt, deren Expression in
10 dem Organismus, verglichen mit dem Wildtyp, einer reduzierten Regulation unterliegt.

- Unter einer reduzierten Regulation verglichen mit dem Wildtyp, wird eine im Vergleich zum vorstehend definierten Wildtyp verringerte, vorzugsweise
15 keine Regulation auf Expressions- oder Proteinebene verstanden.

- Die reduzierte Regulation kann vorzugsweise durch einen im Nukleinsäurekonstrukt mit der kodierenden Sequenz funktionell verknüpften Promotor erreicht werden, der in dem Organismus verglichen
20 mit dem Wildtyp-Promoter einer reduzierten Regulation unterliegt.

- Beispielsweise unterliegen die Promotoren ptef1 aus *Blakeslea trispora* und pgpdA aus *Aspergillus nidulans* nur einer reduzierten Regulation und sind daher insbesondere als Promotoren bevorzugt.

25

Diese Promotoren zeigen eine annähernd konstitutive Expression in *Blakeslea trispora*, so dass die transkriptionelle Regulation nicht mehr über die Intermediate der Carotinoidbiosynthese abläuft.

- 30 Die reduzierte Regulation kann in einer weiteren bevorzugten Ausführungsform dadurch erreicht werden, dass man als Nukleinsäure

codierend eine HMG-CoA-Reduktase eine Nukleinsäure verwendet, deren Expression in dem Organismus, verglichen mit der Organismus eigenen, orthologen Nukleinsäure, einer reduzierten Regulation unterliegt.

- 5 Besonders bevorzugt ist die Verwendung einer Nukleinsäure, die nur den katalytischen Bereich der HMG-CoA-Reduktase kodiert (trunkierte (t-)HMG-CoA-Reduktase). Die für die Regulation verantwortliche Membran-Domäne fehlt. Die verwendete Nukleinsäure unterliegt somit einer reduzierten Regulation und führt zu einer Erhöhung der Genexpression
- 10 der HMG-CoA-Reduktase.
- In einer besonders bevorzugten Ausführungsform bringt man Nukleinsäuren in *Blakeslea trispora* ein, welche die Sequenz SEQ ID. NO. 75 enthalten.
- 15 Weitere Beispiele für HMG-CoA-Reduktasen und damit auch für die auf den katalytischen Bereich reduzierten t-HMG-CoA-Reduktasen bzw. die kodierenden Gene lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Homologievergleiche der Sequenzen aus Datenbanken mit der SEQ ID.
- 20 NO. 75 leicht auffinden.

- Weitere Beispiele für HMG-CoA-Reduktasen und damit auch für die auf den katalytischen Bereich reduzierten t-HMG-CoA-Reduktasen bzw. die kodierenden Gene lassen sich weiterhin beispielsweise ausgehend von
- 25 der Sequenz SEQ ID. NO. 75 aus verschiedenen Organismen deren genomische Sequenz nicht bekannt ist, durch Hybridisierungs- und PCR-Techniken in an sich bekannter Weise leicht auffinden.

- In einer besonders bevorzugten Ausführungsform wird die reduzierte
- 30 Regulation dadurch erreicht, dass man als Nukleinsäure codierend eine HMG-CoA-Reduktase eine Nukleinsäure verwendet, deren Expression in

dem Organismus, verglichen mit der Organismus eigenen, orthologen Nukleinsäure, einer reduzierten Regulation unterliegt und einen Promotor verwendet, der in dem Organismus, verglichen mit dem Wildtyp-Promoter einer reduzierten Regulation unterliegt.

5

Entsprechend wird in einer bevorzugten Variante der Erfindung durch die Transformation die Genexpression der Phytoendesaturase ausgeschaltet, so dass das von den Organismen produzierte Phytoen gewonnen werden kann. Der in der Transformation (i) eingesetzte Vektor umfasst daher in
10 in einer Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für ein Fragment des Gens der Phytoendesaturase, insbesondere carB aus *Blakeslea trispora* mit der SEQ ID NO: 69.

Entsprechend wird in einer bevorzugten Variante der Erfindung durch
15 Transformation die Genexpression der Lycopincyclase ausgeschaltet, so dass das von den Organismen produzierte Lycopin gewonnen werden kann. Der in der Transformation eingesetzte Vektor umfasst daher in einer Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für ein Fragment des Gens der Lycopincyclase, insbesondere
20 carR aus *Blakeslea trispora* s. (WO 03/027293).

In einer weiteren bevorzugten Ausführungsform werden die Organismen der Gattung *Blakeslea* beispielsweise in die Lage versetzt Xanthophylle, wie beispielsweise Zeaxanthin oder Astaxanthin herzustellen, indem
25 genetisch veränderten Organismen der Gattung *Blakeslea* im Vergleich zum Wildtyp eine Hydroxylase-Aktivität und/oder eine Ketolase-Aktivität besitzen.

Der in der Transformation (i) eingesetzte Vektor enthält also in einer
30 weiteren, bevorzugten Variante der Erfindung genetische Informationen, die nach Expression eine Ketolase- und/oder Hydroxylase-Aktivität

entfalten, so dass die Organismen Zeaxanthin oder Astaxanthin produzieren.

Unter Ketolase-Aktivität wird die Enzymaktivität einer Ketolase
5 verstanden.

Unter einer Ketolase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -Ionon-Ring von Carotinoiden eine Keto-Gruppe einzuführen.

10 Insbesondere wird unter einer Ketolase ein Protein verstanden, das die enzymatische Aktivität aufweist, β -Carotin in Canthaxanthin umzuwandeln.

15 Dementsprechend wird unter Ketolase-Aktivität die in einer bestimmten Zeit durch das Protein Ketolase umgesetzte Menge β -Carotin bzw. gebildete Menge Canthaxanthin verstanden.

Unter dem Begriff "Wildtyp" wird erfindungsgemäß der entsprechende
20 nicht genetisch veränderte Ausgangsorganismus der Gattung Blakesleaa verstanden.

Je nach Zusammenhang kann unter dem Begriff "Organismus" der Ausgangsorganismus (Wildtyp) der Gattung Blakesleaa oder ein
25 erfindungsgemäßer, genetisch veränderter Organismus der Gattung Blakesleaa oder beides verstanden werden.

Vorzugsweise wird unter "Wildtyp" für die Verursachung der Ketolase-Aktivität und für die Verursachung der Hydroxylase-Aktivität jeweils ein
30 Referenzorganismus verstanden.

Dieser Referenzorganismus der Gattung *Blakeslea* ist *Blakeslea trispora* ATCC 14271 oder ATCC 14272, die sich lediglich im Paarungstyp unterscheiden.

- 5 Die Bestimmung der Ketolase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen der Gattung *Blakeslea* und in Wildtyp- bzw. Referenzorganismen erfolgt vorzugsweise unter folgenden Bedingungen:

- Die Bestimmung der Ketolase-Aktivität in Organismen der Gattung *Blakeslea* erfolgt in Anlehnung an die Methode von Fraser et al., (J. Biol. Chem. 272(10): 6128-6135, 1997). Die Ketolase-Aktivität in Extrakten wird mit den Substraten beta-Carotin und Canthaxanthin in Gegenwart von Lipid (Sojalecithin) und Detergens (Natriumchoolat) bestimmt. Substrat/Produkt-Verhältnisse aus den Ketolase-Assays werden mittels
- 15 HPLC ermittelt.

- Der erfindungsgemäße genetisch veränderte Organismus der Gattung *Blakeslea* weist in dieser, bevorzugten Ausführungsform im Vergleich zum genetisch nicht veränderten Wildtyp eine Ketolase-Aktivität auf und ist
- 20 somit vorzugsweise in der Lage, transgen eine Ketolase zu exprimieren.

- In einer weiter bevorzugten Ausführungsform erfolgt die Verursachung der Ketolase-Aktivität in den Organismen der Gattung *Blakeslea* durch Genexpression einer Nukleinsäure kodierend eine Ketolase.
- 25

In dieser bevorzugten Ausführungsform erfolgt die Verursachung der Genexpression einer Nukleinsäure kodierend eine Ketolase vorzugsweise durch Einbringen von Nukleinsäuren, die Ketolasen kodieren in die Ausgangsorganismus der Gattung *Blakeslea*.

Dazu kann prinzipiell jedes Ketolase-Gen, also jede Nukleinsäure die eine Ketolase codiert verwendet werden.

Alle in der Beschreibung erwähnten Nukleinsäuren können beispielsweise
5 eine RNA-, DNA- oder cDNA-Sequenz sein.

Bei genomischen Ketolase-Sequenzen aus eukaryontischen Quellen, die Introns enthalten, sind für den Fall das der Wirtsorganismus der Gattung *Blakesleaa* nicht in der Lage ist oder nicht in die Lage versetzt werden
10 kann, die entsprechenden Ketolase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs zu verwenden.

Beispiele für Nukleinsäuren, kodierend eine Ketolase und die
15 entsprechenden Ketolasen, die im erfindungsgemäßen Verfahren verwendet werden können sind beispielsweise Sequenzen aus:

Haematoccus pluvialis, insbesondere aus *Haematoccus pluvialis* Flotow
em. Wille (Accession NO: X86782; Nukleinsäure: SEQ ID NO: 11, Protein
20 SEQ ID NO: 12),

Haematoccus pluvialis, NIES-144 (Accession NO: D45881; Nukleinsäure:
SEQ ID NO: 13, Protein SEQ ID NO: 14),

25 *Agrobacterium aurantiacum* (Accession NO: D58420; Nukleinsäure: SEQ ID NO: 15, Protein SEQ ID NO: 16),

Alicicalgenes spec. (Accession NO: D58422; Nukleinsäure: SEQ ID NO: 17,
Protein SEQ ID NO: 18),

Paracoccus marcusii (Accession NO: Y15112; Nukleinsäure: SEQ ID NO: 19, Protein SEQ ID NO: 20),

Synechocystis sp. Strain PC6803 (Accession NO: NP442491;
5 Nukleinsäure: SEQ ID NO: 21, Protein SEQ ID NO: 22),

Bradyrhizobium sp. (Accession NO: AF218415; Nukleinsäure: SEQ ID NO: 23, Protein SEQ ID NO: 24),

10 Nostoc sp. Strain PCC7120 (Accession NO: AP003592, BAB74888;
Nukleinsäure: SEQ ID NO: 25, Protein SEQ ID NO: 26),

Nostoc punctiforme ATTC 29133, Nukleinsäure: Acc.-No.
NZ_AABC01000195, Basenpaar 55,604 bis 55,392 (SEQ ID NO: 27);
15 Protein: Acc.-No. ZP_00111258 (SEQ ID NO: 28) (als putatives Protein
annotiert) oder

Nostoc punctiforme ATTC 29133, Nukleinsäure: Acc.-No.
NZ_AABC01000196, Basenpaar 140,571 bis 139,810 (SEQ ID NO: 29),
20 Protein: (SEQ ID NO: 30) (nicht annotiert).

Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene, die im
erfindungsgemäßen Verfahren verwendet werden können, lassen sich
beispielsweise aus verschiedenen Organismen, deren genomische
25 Sequenz bekannt ist, durch Identitätsvergleiche der
Aminosäuresequenzen oder der entsprechenden rückübersetzten
Nukleinsäuresequenzen aus Datenbanken mit den vorstehend
beschriebenen Sequenzen und insbesondere mit den Sequenzen
SEQ ID NO: 12 und/oder 26 und/oder 30 leicht auffinden.

Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 12 und/oder 26 und/oder 30 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

Die Hybridisierung kann unter moderaten (geringe Stringenz) oder vorzugsweise unter stringenten (hohe Stringenz) Bedingungen erfolgen.

10

Solche Hybridisierungsbedingungen sind beispielsweise bei Sambrook, J., Fritsch, E.F., Maniatis, T., in: Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57 oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6 beschrieben.

15

Beispielhaft können die Bedingungen während des Waschschruttes ausgewählt sein aus dem Bereich von Bedingungen begrenzt von solchen mit geringer Stringenz (mit 2X SSC bei 50°C) und solchen mit hoher Stringenz (mit 0,2X SSC bei 50°C, bevorzugt bei 65°C) (20X SSC: 0,3 M Natriumcitrat, 3 M Natriumchlorid, pH 7.0).

20

Darüberhinaus kann die Temperatur während des Waschschruttes von moderaten Bedingungen bei Raumtemperatur, 22°C, bis zu stringenten Bedingungen bei 65°C angehoben werden.

25

Beide Parameter, Salzkonzentration und Temperatur, können gleichzeitig variiert werden, auch kann einer der beiden Parameter konstant gehalten und nur der andere variiert werden. Während der Hybridisierung können auch denaturierende Agenzien wie zum Beispiel Formamid oder SDS

30

eingesetzt werden. In Gegenwart von 50 % Formamid wird die Hybridisierung bevorzugt bei 42°C ausgeführt.

Einige beispielhafte Bedingungen für Hybridisierung und Waschschrift sind

5 infolge gegeben:

- (1) Hybridisierungsbedingungen mit zum Beispiel
 - (i) 4X SSC bei 65°C, oder
 - (ii) 6X SSC bei 45°C, oder
 - 10 (iii) 6X SSC bei 68°C, 100 mg/ml denaturierter Fischsperma-DNA, oder
 - (iv) 6X SSC, 0,5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachssperma-DNA bei 68°C, oder
 - (v) 6XSSC, 0,5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachssperma-DNA, 50 % Formamid bei 42°C, oder
 - 15 (vi) 50 % Formamid, 4X SSC bei 42°C, oder
 - (vii) 50 % (vol/vol) Formamid, 0,1 % Rinderserumalbumin, 0,1 % Ficoll, 0,1 % Polyvinylpyrrolidon, 50 mM Natriumphosphatpuffer pH 6,5, 750 mM NaCl, 75 mM Natriumcitrat bei 42°C, oder
 - (viii) 2X oder 4X SSC bei 50°C (moderate Bedingungen), oder
 - 20 (ix) 30 bis 40 % Formamid, 2X oder 4X SSC bei 42°C (moderate Bedingungen).

- (2) Waschschriffe für jeweils 10 Minuten mit zum Beispiel
 - (i) 0,015 M NaCl/0,0015 M Natriumcitrat/0,1 % SDS bei 50°C, oder
 - 25 (ii) 0,1X SSC bei 65°C, oder
 - (iii) 0,1X SSC, 0,5 % SDS bei 68°C, oder
 - (iv) 0,1X SSC, 0,5 % SDS, 50 % Formamid bei 42°C, oder
 - (v) 0,2X SSC, 0,1 % SDS bei 42°C, oder
 - (vi) 2X SSC bei 65°C (moderate Bedingungen).

In einer bevorzugten Ausführungsform der erfindungsgemäßen genetisch veränderten Organismen der Gattung *Blakeslea* bringt man Nukleinsäuren ein, die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 12 oder eine von dieser Sequenz durch Substitution, Insertion oder

5 Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30 %, bevorzugter mindestens 40 %, bevorzugter mindestens 50 %, bevorzugter mindestens 60 %, bevorzugter mindestens 70 %, bevorzugter mindestens 80 %, besonders bevorzugt mindestens 90 %, insbesondere 91%, 92%, 93%,

10 94%, 95%, 96%, 97%, 98%, 99% auf Aminosäureebene mit der Sequenz z SEQ ID NO: 12 und die enzymatische Eigenschaft einer Ketolase aufweist.

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die wie

15 vorstehend beschrieben durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 12 durch künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

20 In einer weiteren, bevorzugten Ausführungsform der erfindungsgemäßen Verfahren bringt man Nukleinsäuren ein die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 26 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren

25 abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30 %, bevorzugter mindestens 40 %, bevorzugter mindestens 50 %, bevorzugter mindestens 60 %, bevorzugter mindestens 70 %, bevorzugter mindestens 80 %, besonders bevorzugt mindestens 90 % insbesondere 91%, 92%, 93%, 94%, 95%, 96%, 97%,

30 98%, 99% auf Aminosäureebene mit der Sequenz SEQ ID NO: 26 und die enzymatische Eigenschaft einer Ketolase aufweist.

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die, wie vorstehend beschrieben, durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche
5 Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 26 durch künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

In einer weiteren, bevorzugten Ausführungsform der erfindungsgemäßen
10 Verfahren bringt man Nukleinsäuren ein die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 30 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30 %, bevorzugter mindestens 40 %, 15 bevorzugter mindestens 50 %, bevorzugter mindestens 60 %, bevorzugter mindestens 70 %, bevorzugter mindestens 80 %, besonders bevorzugt mindestens 90 %, insbesondere 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% auf Aminosäureebene mit der Sequenz SEQ ID NO: 30 und die enzymatische Eigenschaft einer Ketolase aufweist.

20 Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die, wie vorstehend beschrieben, durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 30 durch
25 künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

Unter dem Begriff "Substitution" ist in der Beschreibung der Austausch einer oder mehrerer Aminosäuren durch eine oder mehrere Aminosäuren
30 zu verstehen. Bevorzugt werden sog. konservative Austausche durchgeführt, bei denen die ersetzte Aminosäure eine ähnliche

Eigenschaft hat wie die ursprüngliche Aminosäure, beispielsweise Austausch von Glu durch Asp, Gln durch Asn, Val durch Ile, Leu durch Ile, Ser durch Thr.

- 5 Deletion ist das Ersetzen einer Aminosäure durch eine direkte Bindung. Bevorzugte Positionen für Deletionen sind die Termini des Polypeptids und die Verknüpfungen zwischen den einzelnen Proteindomänen.

10 Insertionen sind Einfügungen von Aminosäuren in die Polypeptidkette, wobei formal eine direkte Bindung durch ein oder mehrere Aminosäuren ersetzt wird.

Unter Identität zwischen zwei Proteinen wird die Identität der Aminosäuren über die jeweils gesamte Proteinlänge verstanden, insbesondere die
15 Identität die durch Vergleich mit Hilfe der Lasergene Software der Firma DNASTAR, Inc. Madison, Wisconsin (USA) unter Anwendung der Clustal Methode (Higgins DG, Sharp PM. Fast and sensitive multiple sequence alignments on a microcomputer. Comput Appl. Biosci. 1989 Apr;5(2):151-1) unter Einstellung folgender Parameter berechnet wird:

20

Multiple alignment parameter:

Gap penalty 10

Gap length penalty 10

Pairwise alignment parameter:

25

K-tuple 1

Gap penalty 3

Window 5

Diagonals saved 5

- 30 Unter einem Protein, das eine Identität von mindestens 20 % auf Aminosäureebene mit der Sequenz SEQ ID NO: 12 oder 26 oder 30

aufweist, wird dementsprechend ein Protein verstanden, das bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 12 oder 26 oder 30, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 20 %, bevorzugt 80 %, 85 %, besonders 90 %, insbesondere 95 % aufweist.

Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

Bevorzugt werden dafür solche Codons verwendet, die entsprechend der Blakesleaspezifischen codon usage häufig verwendet werden. Die codon usage lässt sich anhand von Computerauswertungen anderer, bekannter Gene von Organismen der Gattung Blakeslea leicht ermitteln.

In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 11 in die Organismus der Gattung ein.

In einer weiteren, besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 25 in die Organismus der Gattung ein.

In einer weiteren, besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 29 in die Organismus der Gattung ein.

Alle vorstehend erwähnten Ketolase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die

- chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des
- 5 Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.
- 10 Der in der Transformation (i) eingesetzte Vector umfasst daher in einer Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für eine Ketolase, insbesondere der Ketolase *Nostoc punctiforme* aus mit der SEQ ID NO: 72.
- 15 Unter Hydroxylase-Aktivität die Enzymaktivität einer Hydroxylase verstanden.
- Unter einer Hydroxylase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -
- 20 lonon-Ring von Carotinoiden eine Hydroxy-Gruppe einzuführen.
- Insbesondere wird unter einer Hydroxylase ein Protein verstanden, das die enzymatische Aktivität aufweist, β -Carotin in Zeaxanthin oder Cantaxanthin in Astaxanthin umzuwandeln.
- 25 Dementsprechend wird unter Hydroxylase-Aktivität die in einer bestimmten Zeit durch das Protein Hydroxylase umgesetzte Menge β -Carotin oder Cantaxanthin bzw. gebildete Menge Zeaxanthin oder Astaxanthin verstanden.

Bei einer erhöhten Hydroxylase-Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein Hydroxylase die umgesetzte Menge β -Carotin oder Canthaxanthin bzw. die gebildete Menge Zeaxanthin oder Astaxanthin erhöht.

5

Vorzugsweise beträgt diese Erhöhung der Hydroxylase-Aktivität mindestens 5 %, weiter bevorzugt mindestens 20 %, weiter bevorzugt mindestens 50 %, weiter bevorzugt mindestens 100 %, bevorzugt mindestens 300 %, noch bevorzugt mindestens 500 %, insbesondere

10

mindestens 600 % der Hydroxylase-Aktivität des Wildtyps.

Die Bestimmung der Hydroxylase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen und in Wildtyp- bzw. Referenz-Organismen erfolgt vorzugsweise unter folgenden Bedingungen:

15

Die Aktivität der Hydroxylase wird nach Bouvier et al. (Biochim. Biophys. Acta 1391 (1998), 320-328) *in vitro* bestimmt. Es wird zu einer bestimmten Menge an Organismenextrakt Ferredoxin, Ferredoxin-NADP Oxidoreductase, Katalase, NADPH sowie β -Carotin mit Mono- und

20

Digalaktosylglyzeriden zugegeben.

Besonders bevorzugt erfolgt die Bestimmung der Hydroxylase-Aktivität unter folgenden Bedingungen nach Bouvier, Keller, d'Harlingue und Camara (Xanthophyll biosynthesis: molecular and functional

25

characterization of carotenoid hydroxylases from pepper fruits (*Capsicum annum* L.; Biochim. Biophys. Acta 1391 (1998), 320-328):

Der *in-vitro* Assay wird in einem Volumen von 0,250 ml Volumen durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7,6), 0,025

30

mg Ferredoxin von Spinat, 0,5 Einheiten Ferredoxin-NADP+ Oxidoreduktase von Spinat, 0,25 mM NADPH, 0,010 mg β -Carotin (in

- 0,1 mg Tween 80 emulgiert), 0,05 mM einer Mischung von Mono- und Digalaktosylglyzeriden (1:1), 1 Einheit Katalyse, 200 Mono- und Digalaktosylglyzeriden, (1:1), 0,2 mg Rinderserumalbumin und Organismenextrakt in unterschiedlichem Volumen. Die
- 5 Reaktionsmischung wird 2 Stunden bei 30°C inkubiert. Die Reaktionsprodukte werden mit organischem Lösungsmittel wie THF, Aceton oder Chloroform/Methanol (2:1) extrahiert und mittels HPLC bestimmt.
- 10 Besonders bevorzugt erfolgt die Bestimmung der Hydroxylase-Aktivität unter folgenden Bedingungen nach Bouvier, d'Harlingue und Camara (Molecular Analysis of carotenoid cyclase inhibition; Arch. Biochem. Biophys. 346(1) (1997) 53-64):
- 15 Der in-vitro Assay wird in einem Volumen von 250 µl Volumen durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7,6), unterschiedliche Mengen an Organismenextrakt, 20 nM Lycopin, 250 µg an chromoplastidärem Stromaprotein aus Paprika, 0,2 mM NADP⁺, 0,2 mM NADPH und 1 mM ATP. NADP/NADPH und ATP werden
- 20 in 10 ml Ethanol mit 1 mg Tween 80 unmittelbar vor der Zugabe zum Inkubationsmedium gelöst. Nach einer Reaktionszeit von 60 Minuten bei 30°C wird die Reaktion durch Zugabe von Chloroform/Methanol (2:1) beendet. Die in Chloroform extrahierten Reaktionsprodukte werden mittels HPLC analysiert.
- 25 Ein alternativer Assay mit radioaktivem Substrat ist beschrieben in Fraser und Sandmann (Biochem. Biophys. Res. Comm. 185(1) (1992) 9-15).
- Die Erhöhung der Hydroxylase-Aktivität kann durch verschiedene Wege
- 30 erfolgen, beispielsweise durch Ausschalten von hemmenden Regulationsmechanismen auf Expressions- und Proteinebene oder durch

Erhöhung der Genexpression von Nukleinsäuren kodierend eine Hydroxylase gegenüber dem Wildtyp.

Die Erhöhung der Genexpression der Nukleinsäuren kodierend eine Hydroxylase gegenüber dem Wildtyp kann ebenfalls durch verschiedene
5 Wege erfolgen, beispielsweise durch Induzierung des Hydroxylase-Gens durch Aktivatoren oder durch Einbringen von einer oder mehrerer Hydroxylase-Genkopien, also durch Einbringen mindestens einer Nukleinsäure kodierend eine Hydroxylase in denb Organismus der
10 Gattung *Blakesleaa*.

In einer bevorzugten Ausführungsform erfolgt die Erhöhung der Genexpression einer Nukleinsäure kodierend eine Hydroxylase durch Einbringen von mindestens einer Nukleinsäure kodierend eine
15 Hydroxylase in den Organismus der Gattung *Blakesleaa*.

Dazu kann prinzipiell jedes Hydroxylase-Gen, also jede Nukleinsäure, die eine Hydroxylase codiert, verwendet werden.

20 Bei genomischen Hydroxylase-Sequenzen aus eukaryontischen Quellen, die Introns enthalten, sind für den Fall, dass der Wirtsorganismus nicht in der Lage ist oder nicht in die Lage versetzt werden kann, die entsprechende Hydroxylase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs zu
25 verwenden.

Ein Beispiel für ein Hydroxylase-Gen ist eine Nukleinsäure, codierend eine Hydroxylase aus *Haematococcus pluvialis* mit der Accession No. AX038729 (WO 0061764; Nukleinsäure: SEQ ID NO: 31, Protein: SEQ ID
30 NO: 32), aus *Erwinia uredovora* 20D3 (ATCC 19321, Accession No. D90087; Nukleinsäure: SEQ ID NO: 33, Protein: SEQ ID NO: 34) oder

Hydroxylase aus *Thermus thermophilus* (DE 102 34 126.5) kodiert durch die Sequenz mit der SEQ ID NO 76.

Weitere Hydroxylasen werden von den Nukleinsäuren mit den folgenden

- 5 Accession Nummern kodiert
- emb|CAB55626.1, CAA70427.1, CAA70888.1, CAB55625.1, AF499108_1, AF315289_1, AF296158_1, AAC49443.1, NP_194300.1, NP_200070.1, AAG10430.1, CAC06712.1, AAM88619.1, CAC95130.1, AAL80006.1, AF162276_1, AAO53295.1, AAN85601.1, CRTZ_ERWHE,
- 10 CRTZ_PANAN, BAB79605.1, CRTZ_ALCSP, CRTZ_AGRAU, CAB56060.1, ZP_00094836.1, AAC44852.1, BAC77670.1, NP_745389.1, NP_344225.1, NP_849490.1, ZP_00087019.1, NP_503072.1, NP_852012.1, NP_115929.1, ZP_00013255.1

- 15 In den erfindungsgemäßen bevorzugten transgenen Organismen der Gattung *Blakeslea* liegt also in dieser bevorzugten Ausführungsform gegenüber dem Wildtyp mindestens ein Hydroxylase-Gen vor.

- In dieser bevorzugten Ausführungsform weist der genetisch veränderte
- 20 Organismus beispielsweise mindestens eine exogene Nukleinsäure, kodierend eine Hydroxylase auf.

- Bevorzugt verwendet man in vorstehend beschriebener bevorzugter Ausführungsform als Hydroxylase-Gene Nukleinsäuren, die Proteine
- 25 kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 32, 34 oder kodiert durch die Sequenz mit der SEQ ID NO 76 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 30 %, vorzugsweise mindestens 50 %, bevorzugter mindestens 70%, noch
- 30 bevorzugter mindestens 80 %, am bevorzugtesten mindestens 90%, insbesondere 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% auf

Aminosäureebene mit der Sequenz SEQ. ID. NO: 32, 34 oder kodiert durch die Sequenz mit der SEQ ID NO 76 und die die enzymatische Eigenschaft einer Hydroxylase aufweisen.

- 5 Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, wie vorstehend beschrieben, durch Homologievergleiche der Aminosäuresequenzen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit der SEQ ID. NO: 31, 33 oder 76 leicht auffinden.

- Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich weiterhin beispielsweise ausgehend von der Sequenz SEQ ID NO: 31, 33 oder 76 aus verschiedenen Organismen deren genomische Sequenz nicht
15 bekannt ist, wie vorstehend beschrieben, durch Hybridisierungs- und PCR-Techniken in an sich bekannter Weise leicht auffinden.

- In einer weiter besonders bevorzugten Ausführungsform werden zur Erhöhung der Hydroxylase-Aktivität Nukleinsäuren in Organismen
20 eingebracht, die Proteine kodieren, enthaltend die Aminosäuresequenz der Hydroxylase der Sequenz SEQ ID NO: 32, 34 oder kodiert durch die Sequenz mit der SEQ ID NO 76.

- Geeignete Nukleinsäuresequenzen sind beispielsweise durch
25 Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

- Bevorzugt werden dafür solche Codons verwendet, die entsprechend der Organismenspezifischen codon usage häufig verwendet werden. Die
30 codon usage lässt sich anhand von Computerauswertungen anderer, bekannter Gene der betreffenden Organismen leicht ermitteln.

In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ. ID. NO: 31, 33 oder 76 in den Organismus ein.

5

Alle vorstehend erwähnten Hydroxylase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappenden, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in
10 bekannter Weise, nach der Phosphoamiditmethode (Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie
15 allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

Der in der Transformation (i) eingesetzte Vector umfasst daher in weiteren Ausführungsformen der Erfindung bevorzugterweise eine Sequenz codierend für eine Hydroxylase, insbesondere eine Hydroxylase aus
20 *Haematococcus pluvialis* mit der SEQ ID NO: 70 oder eine Hydroxylase aus *Erwinia uredovae* mit der SEQ ID NO: 71. oder eine Hydroxylase aus *Thermus thermophilus* kodiert durch die Sequenz mit der SEQ ID NO 76.

25

Der in der Transformation (i) eingesetzte Vector enthält vorzugsweise ferner die Expression regelnde und unterstützende Bereiche, insbesondere Promotoren und Terminatoren.

30

Der in der Transformation (i) eingesetzte Vector enthält vorzugsweise den *gpd* und/oder den *ptef1* Promotor und/oder den *trpC* Terminator. Diese

haben sich zur Transformation der *Blakeslea* besonders bewährt. Auch der Einsatz von dem Fachmann geläufigen "inverted repeats" (IR, Römp Lexikon der Biotechnologie 1992, Thieme Verlag Stuttgart, Seite 407 "Invers repetitive Sequenzen") zur Regelung der Expression bzw.

5 Transkription liegt im Rahmen der Erfindung.

Vorteilhafterweise weist der im Vector eingesetzte *gpd* Promotor die Sequenz SEQ ID NO: 1 auf. Vorteilhafterweise weist der im Vector eingesetzte *trpC* Terminator die Sequenz SEQ ID NO: 2 auf.

10 Vorteilhafterweise weist der im Vector eingesetzte *ptef1* Promotor die Sequenz SEQ ID NO: 35 auf.

Insbesondere werden dabei der *gpd* Promotor und der *trpC* Terminator aus *Aspergillus nidulans* und der *ptef1* Promotor aus *Blakeslea trispora*

15 eingesetzt.

Insbesondere enthält der in der Transformation (i) eingesetzte Vector ein Resistenzgen. Bevorzugterweise handelt es sich um ein Hygromycin-Resistenzgen (*hph*), insbesondere das aus *E. coli*. Dieses Resistenzgen

20 hat sich bei dem Nachweis der Transformation und Selektion der Zellen als besonders geeignet herausgestellt.

Als Promotor für *hph* wird also bevorzugt *p-gpdA*, der Promotor der Glycerinaldehyd-3-phosphatdehydrogenase aus *Aspergillus nidulans*

25 genutzt. Als Terminator für *hph* wird bevorzugt *t-trpC*, der Terminator des Gens *trpC*, codierend für Anthranilatsynthasekomponenten aus *Aspergillus nidulans* genutzt.

Als Vektoren haben sich Abkömmlinge des *pBinAHy* Vectors als

30 besonders geeignet herausgestellt. Der zur Transformation eingesetzte Vector umfasst also bevorzugterweise die SEQ ID NO: 3.

Hinzu kommen je nach gewünschtem Carotinoid oder dessen Vorstufe eine Sequenz codierend für eine Hydroxylase, Ketolase, Phytoendesaturase usw. wie diese zuvor beschrieben wurden. Die Vektoren umfassen also in einer Ausführungsform der Erfindung die

5 Sequenz SEQ ID NO: 69 codierend für die Phytoendesaturase. Die Vektoren umfassen ferner in einer weiteren Ausführungsform der Erfindung die Sequenz SEQ ID NO: 72 codierend für eine Ketolase. Die Vektoren umfassen weiter in einer weiteren Ausführungsform der Erfindung die Sequenz SEQ ID NO: 70 oder 71 oder 76 codierend für eine Hydroxylase.

10 Entsprechende Kombinationen der zuvorgenannten Sequenzen liegen ebenso im Rahmen der Erfindung. So umfasst der Vector in einer Ausführungsform sowohl eine Sequenz SEQ ID NO: 72 codierend für eine Ketolase als auch die Sequenz SEQ ID NO: 70 oder 71 oder 76 codierend für eine Hydroxylase und ermöglicht so die Herstellung von Astaxanthin.

15

Insbesondere sind Vektoren ausgewählt aus der Gruppe bestehend aus den SEQ ID NO: 37 bis 51 und 62 im Rahmen der Erfindung einsetzbar.

Mit dem erfindungsgemäßen Verfahren sind gentechnisch veränderte

20 Organismen *Blakeslea*, insbesondere der Art *Blakeslea trispora* bzw. aus ihnen gebildetes Mycel erhältlich.

Die genetisch veränderten Organismen können zur Produktion von Carotinoiden, Xanthophyllen oder deren Vorstufen, insbesondere Phytoen,

25 Bixin, Astaxanthin, Zeaxanthin und Canthaxanthin verwendet werden. Auch können neue, im Wildtyp natürlicherweise nicht vorkommende Carotinoide durch Einbringung der entsprechenden genetischen Information von den gezielt genetisch veränderten Zellen bzw. dem durch sie gebildeten Mycel erzeugt und anschließend isoliert werden.

30

Bevorzugterweise ist die Gewinnung von Carotinoiden oder deren Vorstufen mit den gezielt genetisch veränderten Zellen bzw. das durch sie gebildete Mycel möglich.

- 5 Wird die gentechnische Veränderung nur in Zellen eines der vorkommenden Paarungstypen (bei *Blakeslea trispora* (+) oder (-)) durchgeführt, so wird zur Kultivierung der entsprechend andere, nicht veränderte Paarungstyp zugesetzt, da so eine gute Produktion der Carotinoide oder deren Vorstufen aufgrund der von dem zweiten, nicht
10 veränderten Paarungstyp abgegebenen Substanzen (z. B. Trisporensäuren) zu erreichen ist. Vorteilhafterweise wird jedoch die gentechnische Veränderung in Zellen beider Paarungstypen vorgenommen und diese zusammen kultiviert. Hierdurch wird ein besonders gutes Wachstum und eine optimale Produktion der Carotinoiden oder deren Vorstufen erreicht.
15 Auch eine (künstliche) Zugabe der Trisporensäuren ist möglich und sinnvoll.

- Trisporensäuren sind Sexualhormone in Mucorales Pilzen, wie *Blakeslea*, welche die Bildung von Zygothoren und die Produktion von β -Carotin stimulieren (van den Ende 1968, J. Bacteriol. 96:1298 - 1303, Austin et al.
20 1969, Nature 223:1178 - 1179, Reschke Tetrahedron Lett. 29:3435 - 3439, van den Ende 1970, J. Bacteriol. 101:423 - 428).

Die Erfindung wird nachfolgend an Hand von Beispielen näher ausgeführt.

25 **Material und Methoden**

Molekulargenetische Arbeiten wurden, wenn nicht anders beschrieben, nach den Methoden in Current Protocols in Molecular Biology (Ausubel et al., 1999, John Wiley & Sons) durchgeführt.

30 **Stämme und Wachstumsbedingungen**

Die *Blakeslea trispora* Stämme ATCC 14271 (Paarungstyp(+)) und ATCC14272 (Paarungstyp (-)) wurden von der American Type Culture Collection erhalten. Die Anzucht von *B. trispora* erfolgte in MEP-Medium (Malzextrakt-Pepton-Medium): 30 g/l Malzextrakt (Difco), 3 g/l Pepton (Soytone, Difco), 20 g/l Agar, Einstellung pH 5,5, ad 1000 ml mit H₂O bei 28 °C.

Die Anzucht von *Agrobacterium tumefaciens* LBA4404 erfolgte nach Hoekema et al. (1983, Nature 303:179-180) bei 28 °C für 24 h in Agrobakterien-Minimal Medium (AMM): 10 mM K₂HPO₄, 10 mM KH₂PO₄, 10 mM Glucose, MM-Salze (2,5 mM NaCl, 2 mM MgSO₄, 700 µM CaCl₂, 9 µM FeSO₄, 4 mM (NH₄)₂SO₄).

Transformation von *Agrobacterium tumefaciens*

Das Plasmid pBinAHyg wurde in den Agrobakterienstamm LBA 4404 (Hoekema et al., 1983, Nature 303:179-180) elektroporiert (Mozo and Hooykaas, 1991, Plant Mol. Biol. 16:917-918). Zur Selektion wurden bei der Agrobakterienanzucht folgende Antibiotika verwendet: Rifampicin 50 mg/l (Selektion auf das *A. tumefaciens* Chromosom), Streptomycin 30 mg/l (Selektion auf das Helferplasmid) und Kanamycin 100 mg/l (Selektion auf den binären Vektor).

Transformation von *Blakeslea trispora*

Zur Transformation wurden die Agrobakterien nach 24 h Anzucht in AMM auf eine OD₆₀₀ von 0,15 in Induktionsmedium (IM: MM-Salze, 40 mM MES (pH 5,6), 5 mM Glucose, 2 mM Phosphat, 0,5% Glycerol, 200 µM Acetosyringone) verdünnt und erneut über Nacht in IM bis zu einer OD₆₀₀ von ca. 0,6 angezogen.

Zur Co-Inkubation von *Blakeslea* ATCC 14271 bzw. ATCC14272 und *Agrobacterium* wurden 100 µl Agrobakteriensuspension mit 100 µl

Blakeslea Sporensuspension (10^7 Sporen/ml in 0,9% NaCl) gemischt und steril auf einer Nylon Membran (Hybond N, Amersham) auf IM-Agarose Platten (IM + 18 g/l Agar) verteilt. Nach 3 Tagen Inkubation bei 26 °C wurde die Membran auf eine MEP-Agarplatte (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) überführt. Zur Selektion auf transformierte Blakesleazellen enthält das Medium Hygromycin in einer Konzentration von 100 mg/l sowie zur Selektion gegen Agrobakterien 100 mg/l Cefotaxim. Die Inkubation erfolgte für ca. 7 Tage bei 26 °C. Anschließend erfolgte der Transfer von Mycel auf frische Selektionsplatten. Gebildete Sporen wurden mit 0,9% NaCl abgespült und auf CM17-1-Agar (3 g/l Glucose, 200 mg/l L-Asparagin, 50 mg/l $MgSO_4 \times 7H_2O$, 150 mg/l KH_2PO_4 , 25 µg/l ThiaminHCl, 100 mg/l Yeast Extract, 100 mg/l Na-desoxycholat, 100 mg/l Hygromycin, 100 mg/l Cefotaxim, pH 5,5, 18 g/l Agar) ausplattiert. Zur Isolierung einzelner gentechnisch veränderter Sporen wurden die Sporen durch ein FACS Gerät der Fa. BectonDickson (Modell Vantage+Diva Option) einzeln auf Selektivmedium abgelegt.

Herstellung genetisch veränderter Blakeslea trispora durch Agrobacterium-vermittelte Transformation

Herstellung des rekombinanten Plasmids pBinAHyg

Aus dem Plasmid pANsCos1 (Fig.1, Osiewicz, 1994, Curr. Genet. 26:87-90, SEQ ID NO: 4) wurde die *gpdA-hph-trpC*-Kassette als BglIII/HindIII Fragment isoliert und in das mit BamHI/HindIII geöffnete binäre Plasmid pBin19 (Bevan, 1984, Nucleic Acids Res. 12:8711-8721) ligiert. Der so erhaltene Vektor wurde als pBinAHyg bezeichnet (Fig. 2, SEQ ID NO: 3) und enthielt das *E. coli* Hygromycin-Resistenzgen (*hph*) unter Kontrolle des *gpd* Promotors (SEQ ID NO: 1) und des *trpC* Terminators (SEQ ID NO: 2) aus *Aspergillus nidulans* sowie die entsprechenden Bordersequenzen, die für den DNA-Transfer von *Agrobacterium* notwendig sind. Die in den weiter unten beschriebenen

Ausführungsbeispielen genannten Vektoren sind Abkömmlinge von pBinAHyg.

Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in

5 Agrobacterium tumefaciens

Nachfolgend wird beispielhaft die Übertragung des Plasmids pBinAHyg in Agrobakterien beschrieben. Die Übertragung der Abkömmlinge erfolgte analog.

- 10 Das Plasmid pBinAHyg wurde in den Agrobakterienstamm LBA 4404 (Hoekema et al., 1983, Nature 303:179-180) elektroporiert (Mozo and Hooykaas, 1991, Plant Mol. Biol. 16:917-918). Zur Selektion wurden bei der Agrobakterienanzucht folgende Antibiotika verwendet: Rifampicin 50 mg/l (Selektion auf das *A. tumefaciens* Chromosom), Streptomycin 30 mg/l
- 15 (Selektion auf das Helferplasmid) und Kanamycin 100 mg/l (Selektion auf den binären Vektor).

Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in Blakeslea trispora

- 20 Zur Transformation wurden die Agrobakterien nach 24 h Anzucht in AMM auf eine OD₆₆₀ von 0,15 in Induktionsmedium (IM: MM-Salze, 40 mM MES (pH 5,6), 5 mM Glucose, 2 mM Phosphat, 0,5% Glycerol, 200 µM Acetosyringone) verdünnt und erneut über Nacht in IM bis zu einer OD₆₆₀ von ca. 0,6 angezogen.

25

- Zur Co-Inkubation von *Blakeslea trispora* (B.t.) und *Agrobacterium tumefaciens* (A.t.) wurden 100 µl Agrobaktériensuspension mit 100 µl *Blakeslea* Sporensuspension (10⁷ Sporen/ml in 0,9% NaCl) gemischt und steril auf einer Nylon Membran (Hybond N, Amersham) auf IM-Agarose
- 30 Platten (IM + 18 g/l Agar) verteilt. Nach 3 Tagen Inkubation bei 26 °C

wurde die Membran auf eine MEP-Agarplatte (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) überführt.

Zur Selektion auf transformierte *Blakeslea*-Zellen enthielt das Medium Hygromycin in einer Konzentration von 100 mg/l sowie zur Selektion gegen Agrobakterien 100 mg/l Cefotaxim. Die Inkubation erfolgte für ca. 7 Tage bei 26 °C. Anschließend erfolgte der Transfer von Mycel auf frische Selektionsplatten. Gebildete Sporen wurden mit 0,9% NaCl abgespült und auf CM17-1-Agar (3 g/l Glucose, 200 mg/l L-Asparagin, 50 mg/l $\text{MgSO}_4 \times 7\text{H}_2\text{O}$, 150 mg/l KH_2PO_4 , 25 µg/l Thiamin-HCl, 100 mg/l Yeast Extract, 100 mg/l Na-desoxycholat, pH 5,5, 100 mg/l Cefotaxim, 100 mg/l Hygromycin, 18 g/l Agar) ausplattiert. Die Übertragung von Sporen auf frische Selektionsplatten wurde dreimal wiederholt. Auf diese Weise wurde die Transformante *Blakeslea trispora* GVO 3005 isoliert. Alternativ erfolgte zur Selektion der GVO (gentechnisch veränderten Organismen) die Einzelablage der Sporen durch den BectonDickinson FacsVantage+Diva Option auf CM-17 Agar mit 100 mg/l Cefotaxim, 100 mg/l Hygromycin. In diesem Fall wurde nur dort Pilzmycel gebildet, wo die Sporen gentechnisch verändert waren.

Nachweis der genetischen Veränderung durch Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in *Blakeslea trispora*
Nachfolgend wird beispielhaft der Nachweis der Übertragung für pBinAHyg in *Blakeslea trispora* beschrieben. Der Nachweis der Übertragung der Abkömmlinge erfolgte analog.

200 ml MEP-Medium (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5) wurden mit 10^5 bis 10^7 Sporen der Transformante *Blakeslea trispora* GVO 3005 beimpft und 7 Tage bei 26 °C mit 200 Upm auf einem Rundschtüttler inkubiert. Zum Nachweis der erfolgreichen Transformation wurde DNA aus dem Mycel isoliert (Pqclab Fungal DNA Mini Kit) und in einer PCR

(Programm: 94 °C 1 min, dann 30 Zyklen mit 1 min. 94°C, 1 min. 58 °C, 1 min. 72 °C) eingesetzt.

Zum Nachweis des Hygromycinresistenzgens (hph) wurden die Primer
 5 hph-forward (5'-CGATGTAGGAGGGCGTGGATA, SEQ ID NO: 5) und
 hph-reverse (5'-GCTTCTGCGGGCGATTGTGT, SEQ ID NO: 6)
 verwendet. Das erwartete Fragment von hph wies eine Länge von 800 bp
 auf.

10 Zur Amplifikation des Kanamycinresistenzgens nptIII und damit als
 Kontrolle auf Agrobakterien wurden die Primer nptIII-forward (5'-
 TGAGAATATCACCGGAATTG, SEQ ID NO: 7) und nptIII-reverse (5'-
 AGCTCGACATACTGTTCTTCC, SEQ ID NO: 8) verwendet. Das
 erwartete Fragment von nptIII wies eine Länge von 700 bp auf.

15 Zur Amplifikation eines Fragmentes des Glycerinaldehyd-3-
 phosphatdehydrogenasegens gpd1 und damit als Kontrolle auf Blakeslea
 trispora wurden die Primer MAT292 (5'-
 GTGAATGGAAATCCCATCGCTGTC, SEQ ID NO: 9) und MAT293 (5'-
 20 AGTGGGTACTCTAAAGGCCATACC, SEQ ID NO: 10) verwendet. Das
 erwartete Fragment von gpd1 wies eine Länge von 500 bp auf.

Das Ergebnis der PCR der Blakeslea trispora DNA ist in Fig. 3 anhand
 eines Standard-Gels gezeigt. Die Spuren des Gels wurden
 25 folgendermaßen belegt:

- | | | |
|----|--|--------------------------------|
| | 1) 100 bp Größenmarker (100 bp - 1 kb) | |
| | 2) B.t. GVO 3005 | primer nptIII-for / nptIII-rev |
| | 3) B.t. GVO 3005 | primer hph-for / hph-rev |
| 30 | 4) B.t. GVO 3005 | primer MAT292 / MAT293 (gpd) |
| | 5) A.t. mit Plasmid pBinAHyg | primer nptIII-for / nptIII-rev |

- | | |
|------------------------------|--------------------------------|
| 6) A.t. mit Plasmid pBinAHyg | primer hph-for / hph-rev |
| 7) B.t. 14272 WT | primer nptIII-for / nptIII-rev |
| 8) B.t. 14272 WT | primer hph-for / hph-rev |
| 9) B.t. 14272 WT | primer MAT292 / MAT293 (gpd) |

5

In der DNA von *Blakeslea trispora* wurde das Hygromycinresistenzgens (hph) und als Positivkontrolle Glycerinaldehyd-3-phosphatdehydrogenasegen (gpd1) nachgewiesen. nptIII konnte demgegenüber nicht nachgewiesen werden.

10

Somit wurde die genetische Veränderung von *Blakeslea trispora* durch *Agrobacterium*-vermittelte Transformation nachgewiesen.

Isolierung homokaryotischer GVO von *Blakeslea trispora*:

- 15 Durch erfolgreichen Transfer des Vectors pBinAHyg und Abkömmlingen von pBinAHyg in *Blakeslea trispora* entstehen genetisch veränderte Organismen (GVO) von *Blakeslea trispora*. Jedoch liegen in *Blakeslea* in allen Stadien des vegetativen und des sexuellen Zellzyklus mehrkernige Zellen vor. Daher erfolgt die Insertion der Fremd-DNA in der Regel nur in
- 20 einem Kern. Ziel ist es, Stämme von *Blakeslea* zu erhalten, bei denen die Insertion der Fremd-DNA in allen Kernen vorliegt, d.h. Ziel ist ein homonukleates rekombinantes Pilzmycel.

1) Herstellung homonukleater rekombinanter Stämme durch FACS (fluorescence-activated cell sorting)

25

Ein geringer Anteil der Sporen von *Blakeslea trispora* bzw. der gentechnisch veränderten Stämme von *Blakeslea trispora* ist von Natur aus einkernig. Zur Herstellung homonukleater rekombinanter Stämme, die Fremd-DNA von pBinAHyg oder pBinAHyg-Abkömmlingen enthielten,

30 wurden die einkernigen Sporen durch FACS aussortiert und auf MEP (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) mit 100 mg/l Cefotaxim

und 100 mg/l Hygromycin plattiert. Die hier gebildeten Mycelien waren homonukleat. Zur Sortierung mit FACS wurden die Sporen eines 3 Tage alten Ausstriches mit 10 ml Tris-HCl 50mMol + 0,1% Span20 pro Agar-Platte abgeschwemmt. Die Sporenkonzentration betrug 0,5 bis $0,8 \times 10^7$ Sporen pro ml. Zu 9 ml Sporensuspension wurden 1ml DMSO und 10 µl Syto 11 (Farbstoff-Stammlösung In DMSO Molecular Probes Nr.S-7573) zugegeben. Danach wurde 2 h bei 30°C gefärbt. Selektion und Ablage erfolgten mittels eines Gerätes vom Typ FacsVantage+Diva Option Fa. Becton Dickinson. Die Selektion erfolgte zuerst nach Größe, um einzelne Sporen von Aggregaten und Verunreinigungen zu trennen. Dann wurden diese Sporen nach ihrer Fluoreszenz (Anregung = 488nm; Emission = 530 nm) sortiert abgelegt. Die linke Schulter der Gauß-Kurve der Fluoreszenzhäufigkeitsverteilung enthielt die einkernigen Sporen.

2) Herstellung homonukleater Stämme durch Kernreduktion und Selektion mit FACS

Zur Reduzierung der Anzahl von Kernen pro Spore wurde vor der Selektion eine Behandlung von Sporensuspensionen mit MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) durchgeführt, und so durch chemische Mutagenese eine Kernreduktion erzielt.

Hierfür wurde zunächst eine Sporensuspension mit 1×10^7 Sporen/ml in Tris/HCl-Puffer, pH 7,0 hergestellt. Der Sporensuspension wurde MNNG in einer Endkonzentration von 100 µg/ml zugegeben. Die Zeit der Inkubation in MNNG wurde so gewählt, dass die Überlebensrate der Sporen ca. 5% betrug. Nach Inkubation mit MNNG wurden die Sporen dreimal mit 1g/l Span 20 in 50 mM Phosphatpuffer pH 7,0 gewaschen und nach der unter 1) beschriebenen Methode sortiert bzw. selektiert.

Alternativ konnten zur Reduktion der Kernzahl in den Sporen auch Röntgen- und UV-Strahlen eingesetzt werden, wie es von Cerdá-Olmedo und Patricia Reau in *Mutation Res.*, 9 (1970), 369-384 beschrieben wurde.

5 **3) Herstellung homonukleater Stämme durch Selektion auf rezessive Selektionsmarker**

Als rezessiver Selektionsmarker zur Selektion homonukleater Mycelien kommt beispielsweise der rezessive Selektionsmarker *pyrG* in Frage. Wildtyp-Stämme von *Blakeslea trispora* sind *pyrG*⁺. Diese Stämme können nicht in Gegenwart des Pyrimidin-Analogs 5-Fluororotat (FOA) wachsen, weil sie FOA durch die Orotidin-5'-monophosphatdecarboxylase zu lethalen Metaboliten umsetzen. Gentechnisch veränderte *Blakesleae*, die homonukleat *pyrG*⁻ sind, fehlt die Enzymaktivität Orotidin-5'-monophosphatdecarboxylase. Folglich können diese *pyrG*⁻-Stämme 5-Fluororotat nicht verwerten. Die Stämme wachsen daher in Gegenwart von FOA und Uracil. Im Fall der Kopplung der Mutation *pyrG*⁻ und der Insertion von Fremd-DNA auf dem Kern einer einkernigen Spore, kann aus dieser Spore homonukleates rekombinantes Pilzmycel gebildet werden.

20 Zunächst wurde durch Insertion eines Fragmentes von *pyrG* (SEQ ID NO: 65) aus *Blakeslea trispora* in pBinAHyg das Plasmid pBinAHygBTpyrG-SCO (SEQ ID NO: 36, Fig. 4) erzeugt. Dieses Plasmid wurde in *Blakeslea trispora* transformiert und führte dort durch homologe Rekombination zur
25 Disruption von *pyrG*.

Homonukleate GVO von *Blakeslea trispora* mit dem Phänotyp *pyrG*⁻ wurden folgendermaßen selektiert. Zur Agrobakterium-vermittelten Transformation von pBinAHygBTpyrG-SCO wurde wie oben beschrieben
30 auf MEP (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) mit 100 mg/l Cefotaxim und 100 mg/l Hygromycin plattiert. Die Sporen der

Transformanten wurden mit 10 ml Tris-HCl 50mM + 0,1% Span20 pro Agar-Platte abgeschwemmt. Die Sporenkonzentration betrug $0,5 \text{ bis } 0,8 \times 10^7$ Sporen pro ml. Die Sporen wurden anschließend auf FOA-Medium mit 100 mg/l Cefotaxim und 100 mg/l Hygromycin ausplattiert. FOA-Medium
5 enthielt pro Liter 20 g Glucose, 1 g FOA, 50 mg Uracil, 200 ml Citrat-Puffer (0,5 M, pH 4,5) und 40 ml Spurensalzlösung nach Sutter, 1975, PNAS, 72:127). Homonukleate pyrG^- -Mutanten zeigten Wachstum auf dem Uracil-haltigen FOA-Medium; aber kein Wachstum bei Plattierung auf FOA-Medium ohne Uracil. Auf die gleiche Weise wurden aus den im
10 folgenden beschriebenen GVO von *Blakeslea trispora* zur Herstellung von Xanthophyllen homonukleate GVO hergestellt.

Alternativ ist es möglich die Sporen analog zur Vorschrift von Roncero et al. auf Medium mit 5-Carbon-5-deazariboflavin zu plattieren, das
15 zusätzlich Hygromycin enthält (Roncero et al., 1984, Mutation Research, 125: 195 - 204). Hierdurch werden homokaryonte Zellen des Genotyps hyg^R und dar^- selektiert. Nach diesem Prinzip werden homokaryonte Stämme von *Blakeslea trispora* mit dem Phänotyp hyg^R und dar^- erzeugt.

20

Ausführungsbeispiele zur Herstellung von gentechnisch veränderten Organismen von *Blakeslea trispora* für die Herstellung von Carotinoiden und Carotinoidvorstufen

25 Die Erzeugung der im folgenden genannten Plasmide erfolgte durch die Methode „overlap-extension PCR“ und durch anschließende Insertion der Amplifikationsprodukte in das Plasmid pBinAHyg. Die Methode „overlap-extension PCR“ erfolgte wie in Innis et al. (Eds.) PCR protocols: a guide to methods and applications, Academic Press, San Diego beschrieben. Die
30 Transformation der pBinAHyg-Abkömmlinge und die Herstellung

homonukleater gentechnisch veränderter Stämme von *Blakeslea trispora* erfolgte wie oben beschrieben.

Gentechnisch veränderte Stämme von *Blakeslea trispora* zur

5 Herstellung von Zeaxanthin

Folgende Plasmide (Abkömmlinge von pBinAHyg) wurden zur gentechnischen Veränderung von *Blakeslea trispora* für die Herstellung von Zeaxanthin verwendet, codieren also u.a. Hydroxylasen (crtZ):

- 10 - p-tef1-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ (SEQ ID NO: 70) aus *Haematococcus pluvialis* Flotow NIES-144 (Accession No. AF162276) unter Kontrolle des ptef1 Promotors aus *Blakeslea trispora* (Seq. pBinAHygBTpTEF1-HPcrtZ, SEQ ID NO: 37, Fig. 5);
- 15 - p-carRA-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHyg-BTpcarRA-HPcrtZ, SEQ ID NO: 38, Fig. 6)
- 20 - p-carB-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarB aus *Blakeslea trispora* (Seq. pBinAHygBTpcarB-HPcrtZ, SEQ ID NO: 39, Fig. 7)
- 25 - p-carRA-HPcrtZ-TAG-3'carA-IR, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora*. Stromabwärts des Gens der Hydroxylase ist eine Inverted-Repeat-Struktur lokalisiert, die aus dem 3'-Ende von carA und der stromabwärts von carA gelegenen Region stammt (IR, SEQ ID NO: 74, 'Inverted Repeat 1' ca. 350 bp von carA, dann ca. 200 bp 'Loop' und anschließend ca. 350 bp 'Inverted Repeat 2') (Seq. pBinAHyg-BTpcarRA-HPcrtZ-TAG-3'carA-IR, SEQ ID NO: 40, Fig. 8);

- p-carRA-HPcrtZ-GCG-3'carA-IR, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora*. Das Gen der Hydroxylase ist mit einer Inverted-Repeat-Struktur fusioniert, die aus dem 3'-Ende von carA und der stromabwärts von carA gelegenen Region stammt (IR, SEQ ID NO: 74, 'Inverted Repeat 1' ca. 350 bp von carA, dann ca. 200 bp 'Loop' und anschließend ca. 350 bp 'Inverted Repeat 2'). Das abgeleitete Fusionsprotein besteht folglich aus der Hydroxylase von *Haematococcus pluvialis* und dem Carboxyterminus von CarA aus *Blakeslea trispora* (Seq. pBinAHyg-BTp-carRA-HPcrtZ-GCG-3'carA-IR, SEQ ID NO: 41, Fig. 9);
- p-tef1-EUcrtZ, enthaltend Gen der Hydroxylase EUcrtZ (SEQ ID NO: 71) aus *Erwinia uredova* 20D3 (Accession No. D90087) unter Kontrolle des ptef1 Promotors (Seq. pBinAHygBTpTEF1-EUcrtZ, SEQ ID NO: 42, Fig. 10);
- p-carRA-EUcrtZ, enthaltend Gen der Hydroxylase EUcrtZ aus *Erwinia uredova* 20D3 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHygBTp-carRA-EUcrtZ, SEQ ID NO: 43, Fig. 11);
- p-carB-EUcrtZ, enthaltend Gen der Hydroxylase EUcrtZ aus *Erwinia uredova* 20D3 unter Kontrolle des Promotors pcarB aus *Blakeslea trispora* (Seq. pBinAHygBTp-carB-EUcrtZ, SEQ ID NO: 44, Fig. 12);
- p-gpdA-HPcrtZ-t-crtZ, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des gpdA Promotors und des Terminators t-crtZ; d.h. des stromabwärts von crtZ aus *Haematococcus pluvialis* Flotow NIES-144 gelegenen Sequenzabschnitts (SEQ ID NO: 73) (Seq. pBinAHyg-gpdA-HPcrtZ-t-crtZ, SEQ ID NO: 45, Fig. 13).

- p-gpdA-BTcarR-HPcrtZ-BTcarA, enthaltend Genfusion aus Genen der Lycopinecyclase carR aus *Blakeslea trispora*, der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 und der Phytoensynthase carA aus *Blakeslea trispora* unter Kontrolle des gpdA Promotors aus *Aspergillus nidulans* (Seq. pBinAHyg-carR_crtZ_carA, SEQ ID NO: 46, Fig. 14);

Herstellung gentechnisch veränderter Stämme von *Blakeslea trispora* zur Herstellung von Canthaxanthin

10 Folgende Plasmide (Abkömmlinge von pBinAHyg) wurden zur gentechnischen Veränderung von *Blakeslea trispora* für die Herstellung von Canthaxanthin verwendet, codieren also u.a. Ketolasen (crtW):

- p-lef1-NPcrtW, enthaltend das Gen der Ketolase NPcrtW (SEQ ID NO: 72) aus *Nostoc punctiforme* PCC73102 (ORF148, Accession No. NZ_AABC01000196) unter Kontrolle des ptef1 Promotors aus *Blakeslea trispora* (Seq. pBinAHygBTpTEF1-NpucrtW, SEQ ID NO: 47, Fig. 15);
- p-carRA-NPcrtW, enthaltend das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 unter der Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHygBTpcarRA-NpucrtW, SEQ ID NO: 48, Fig. 16);
- p-carB-NPcrtW, enthaltend das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 unter der Kontrolle des Promotors pcarB aus *Blakeslea trispora* (Seq. pBinAHygBTpcarB-NpucrtW, SEQ ID NO: 49, Fig. 17);

Herstellung gentechnisch veränderter Stämme von *Blakeslea trispora* zur Herstellung von Astaxanthin

30 Folgende Plasmide (Abkömmlinge von pBinAHyg) wurden zur gentechnischen Veränderung von *Blakeslea trispora* für die Herstellung

von Astaxanthin verwendet, codieren also u.a. für Hydroxylasen (crtZ) und Ketolasen (crtW):

- p-carRA-HPcrtZ-pcarRA-NPcrtW, enthaltend das Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 und das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 (ORF148, Accession No. NZ_AABC01000196) beide jeweils unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHygBTpcarRA-HPcrtZ-BTpcarRA-NpucrtW, SEQ ID NO: 50, Fig. 18);
- p-carRA-EUcrtZ-pcarRA-NPcrtW, enthaltend das Gen der Hydroxylase EUcrtZ aus *Erwinia uredovae* 20D3 (Accession No. D90087) und das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 beide jeweils unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHygBTpcarRA-EUcrtZ-BTpcarRA-NpucrtW, SEQ ID NO: 51, Fig. 19);

Klonierung und Sequenzanalyse von Genen und Promotoren, die beispielhaft für die gentechnische Veränderung von *Blakeslea trispora* genutzt werden können.

- Nachfolgend werden beispielhaft die Klonierung und Sequenzierung verschiedener Gene und Promotoren aus *Blakeslea trispora* beschrieben.

Klonierung und Sequenzanalyse ptef1

- Die Klonierung von p-tef aus *Blakeslea trispora* erfolgte auf der Grundlage einer bereits in GenBank veröffentlichten Sequenz des Strukturgens für den Translations-Elongationsfaktor 1- α aus *Blakeslea trispora* (AF157235). Ausgehend von dem Sequenzeintrag AF157235 wurden Primer für die inverse PCR ausgewählt, um die stromaufwärts des Strukturgens gelegene Promotoregion zu amplifizieren und zu sequenzieren.

In der inversen nested PCR an 200 ng XhoI-gespaltener und zirkularisierter genomischer DNA von *Blakeslea trispora* ATCC14272 wurde ein 3000-bp-Fragment in folgendem Ansatz erhalten: Matrizen-DNA (1 µg genomische DNA von *Blakeslea trispora* ATCC 14272) Primer
5 MAT344 5'-GGCGTACTTGAAGGAACCCTTACCG-3' (SEQ ID NO: 63) und MAT 345 5'-ATTGATGCTCCCGTCACCGTGATT-3' (SEQ ID NO: 64) je 0,25 µM, 100 µM dNTP, 10 µl Herculanase-Polymerasepuffer 10x, 5 U Herculanase (Zugabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 60 °C, 30 s. 72 °C, 60 s, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus). Der Sequenzabschnitt, der stromaufwärts des vermutlichen Startcodons des Gens *tef1* innerhalb 3000-bp-Fragmentes liegt, wurde als Promotor *ptef1* bezeichnet.

Klonierung Sequenzanalyse des Gens der HMG-CoA-Reduktase aus

15 *Blakeslea trispora*

Zunächst wurde mit dem Cosmidvektor pANsCos1 eine Genbank von *Blakeslea trispora* ATCC 14272, Mating Type (-) hergestellt. Der Vektor wurde durch Spaltung mit XbaI linearisiert und anschließend dephosphoryliert. Eine weitere Spaltung mit BamHI schuf die
20 Insertionsstelle, in welche die mit Sau3AI partiell gespaltene und dephosphorylierte genomische DNA von *Blakeslea trispora* ligiert wurde. Die derart gebildeten Cosmide wurden anschließend *in vitro* verpackt und in *Escherichia coli* übertragen.

Auf der Grundlage der bekannten Sequenz eines Fragmentes des HMG-CoA-Reduktase codierenden Gens aus *Blakeslea trispora* (Eur. J. Biochem 220, 403-408 (1994)) wurde eine 315-bp-DNA-Sonde durch folgende PCR hergestellt. Reaktionsansatz: 1 µg genomische DNA von *Blakeslea trispora* ATCC 14272, Primer MAT314 5'-CCGATGGCGACGACGGAAGTTGTT-3' [SEQ ID NO 79] und MAT315
30 5'-CATGTTTCATGCCATTGCATCACCT-3' [SEQ ID NO 80] je 0,25 µM, 100 µM dNTP, 10 µl Herculanase-Polymerasepuffer 10x, 5 U Herculanase

(Zugabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 58 °C, 30 s, 72 °C, 30 s, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus).

- 5 Mit dieser DNA-Sonde wurde die Cosmid-Genbank durchmustert. Es wurde ein Klon identifiziert, dessen Cosmid mit der DNA-Sonde hybridisierte. Die Insertion dieses Cosmids wurde sequenziert. Die DNA-Sequenz enthielt einen Abschnitt, der dem Gen einer HMG-CoA-Reduktase zugeordnet wurde [SEQ ID NO 75].

10 Klonierung und Sequenzanalyse carB

(carB = Gen der Phytoendesaturase aus *Blakeslea trispora*)

- Aus dem Sequenzvergleich der Peptidsequenzen von Phytoendesaturasen und dem Vergleich der zugehörigen DNA-Sequenzen von *Phycomyces blakesleeanus*, *Cercospora nicotianae*,
15 *Phaffia rhodozyma* und *Neurospora crassa* wurden die degenerierten Primer MAT182 5'-GCNGARGGNATHTGGTA-3' (SEQ ID 52) und MAT192 5'-TCNGCNAGRAADATRTTRTG-3' (SEQ ID 53) abgeleitet. Die PCR wurde in 100 µl Ansätzen durchgeführt. Diese enthielten 200 ng genomische DNA von *Blakeslea trispora* ATCC14272, 1 µM MAT182, 1
20 µM MAT192, 100 µM dNTP, 10 µl Pfu-Polymerasepuffer 10x, 2,5 U Pfu-Polymerase (Zugabe bei 85 °C), H₂O ad 100 µl.

Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 40 °C, 30 s, 72 °C, 30 s, 95 °C, 30 s (35 Zyklen); 72 °C, 10 min (1 Zyklus).

25

- Hiermit wurde ein 358-bp-Fragment erhalten, dessen abgeleitete Peptidsequenz Ähnlichkeit zu den Sequenzen der Phytoendesaturasen aufwies. Durch die Methode der inversen PCR (Innis et al. in PCR protocols: a guide to methods and applications. 1990. S. 219-227) wurden
30 nach dem Prinzip des Chromosome-Walking die Genregionen

stromaufwärts und stromabwärts des 350-bp-Fragmentes folgendermaßen amplifiziert, kloniert und sequenziert:

- (i) ein 1,1-kbp-Fragment durch PCR mit den Primern MAT219 5'-AAGTGACACCGGTTACACGCTTGTCTT-3' (SEQ ID 54) und MAT 220 5'-GCTTATCACCATCTGTTACCTCCTTGC-3' (SEQ ID 55) erhalten aus 200 ng EcoRI-gespaltener und zirkularisierter genomischer DNA von *Blakeslea trispora* ATCC14272, 0,25 µM MAT219, 0,25 µM MAT220, 100 µM dNTP, 10 µl Herculanase-Polymerasepuffer 10x, 5 U Herculanase (Zugabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 60 °C, 30 s, 72 °C, 60 s, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus),
- (ii) ein 2,9-kbp-Fragment durch PCR mit den Primern MAT219 und MAT220 erhalten aus 200 ng XbaI-gespaltener und zirkularisierter genomischer DNA von *Blakeslea trispora* ATCC14272, 0,25 µM MAT219, 0,25 µM MAT220, 100 µM dNTP, 10 µl Herculanase-Polymerasepuffer 10x, 5 U Herculanase (Zugabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 60 °C, 30 s, 72 °C, 3 min, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus);

Der klonierte Sequenzabschnitt ist schematisch in Fig. 20 [SEQ ID NO 77] dargestellt. Die Sequenzierung erfolgte in Strang- und Gegenstrangrichtung mit den klonierten Fragmenten sowie mit den PCR-Produkten. Die Sequenz des klonierten Sequenzabschnitts ist in Fig. 21 [SEQ ID NO 78] gezeigt.

Sequenzvergleiche

Die Nukleotidsequenz von *carB* und die Peptidsequenz des abgeleiteten Proteins CarB wurden mit den bekannten Sequenzen verwandter Proteine verglichen. Zum Sequenzvergleich wurden die Programme GAP und BESTFIT eingesetzt.

CarB - Identische Aminoacylreste nach GAP

Programmeinstellungen:

Gap Weight: 8

5 Length Weight: 2

Average Match: 2.912

Average Mismatch: -2.003

Dabei wurde folgende Werte für die Übereinstimmung der Aminosäuren zu
CarB aus *Blakeslea trispora* ATCC14272 in % gefunden:

10 *Phycomyces blakesleeanus*: 72,491*Phaffia rhodozyma*: 50,460*Neurospora crassa*: 47,943*Cercospora nicotianae*: 47,740

15

CarB -Identische Aminoacylreste nach BESTFIT

Programmeinstellungen:

Gap Weight: 8

Length Weight: 2

20 Average Match: 2.912

Average Mismatch: -2.003

Dabei wurde folgende Werte für die Übereinstimmung der Aminosäuren zu
CarB aus *Blakeslea trispora* ATCC14272 in % gefunden:

Phycomyces blakesleeanus: 73,38025 *Phaffia rhodozyma*: 53,175*Neurospora crassa*: 51,896*Cercospora nicotianae*: 50,791**carB - Identische Basen nach GAP**

30 Programmeinstellungen:

Gap Weight: 50

Length Weight: 3
 Average Match: 10.000
 Average Mismatch: 0.000

Dabei wurde folgende Werte für die Übereinstimmung der Basen zu CarB

5 aus *Blakeslea trispora* ATCC14272 in % gefunden:

Phycomyces blakesleeanus: 64,853
Cercospora nicotianae: 50,143
Phaffia rhodozyma: 43,179
Neurospora crassa: 42,130

10

carB -Identische Basen nach BESTFIT

Programmeinstellungen:

Gap Weight: 50
 Length Weight: 3

15 Average Match: 10.000
 Average Mismatch: -9.000

Dabei wurde folgende Werte für die Übereinstimmung der Basen zu CarB
 aus *Blakeslea trispora* ATCC14272 in % gefunden:

Phycomyces blakesleeanus: 68,926
 20 *Phaffia rhodozyma*: 62,403
Neurospora crassa: 60,230
Cercospora nicotianae: 56,884

Klonierung zur Expression von carB

25 Zur Klonierung und Expression von *carB* aus *Blakeslea trispora* wurden
 von dem oben beschriebenen klonierten Sequenzabschnitt aus *Blakeslea*
trispora in sechs Leserastern die möglichen Proteinsequenzen abgeleitet.
 Diese Proteinsequenzen wurden mit den Sequenzen der
 Phytoendesaturasen aus *Phycomyces blakesleeanus*, *Phaffia rhodozyma*,
 30 *Neurospora crassa*, *Cercospora nicotianae* verglichen. Auf der Grundlage
 des Sequenzvergleiches wurden im klonierten Sequenzabschnitt der

- genomischen DNA von *Blakeslea trispora* drei Exons identifiziert, die zusammengefügt eine codierende Region ergeben, deren abgeleitetes Genprodukt über die gesamte Länge 72,7% identische Aminoacylreste mit der Phytoendesaturase CarB aus *Phycomyces blakesleeanus* aufweist.
- 5 Dieser Sequenzabschnitt aus drei möglichen Exons und zwei möglichen Introns wurde daher als Gen *carB* bezeichnet. Zur Überprüfung der vorhergesagten Genstruktur wurde die codierende Sequenz von *carB* aus *Blakeslea trispora* durch PCR mit cDNA von *Blakeslea trispora* als Matrize und mit den Primern Bol1425 5'-
- 10 AGAGAGGGATCCTTAAATGCGAATATCGTTGC-3' (SEQ ID 56) und Bol1426 5'-AGAGAGGGATCCATGTCTGATCAAAAGAAGCA-3' (SEQ ID 57) erzeugt. Das erhaltene DNA-Fragment wurde sequenziert. Die Lokalisation von Exons und Introns wurde durch Vergleich der cDNA mit der genomischen DNA von *carB* bestätigt. In Fig. 21 ist die codierende
- 15 Sequenz von *carB* schematisch dargestellt. Zur Expression von *carB* in *Escherichia coli* wurde zunächst die NdeI-Schnittstelle in *carB* durch die Methode overlap extension PCR entfernt sowie am 5'-Ende des Gens eine NdeI-Schnittstelle und am 3'-Ende eine BamHI-Schnittstelle eingefügt. Das erhaltene DNA-Fragment wurde mit dem Vektor pJOE2702 ligiert.
- 20 Das erhaltene Plasmid wurde als pBT4 bezeichnet und zusammen mit pCAR-AE in *Escherichia coli* XL1-Blue kloniert. Die Expression erfolgte durch Induktion mit Rhamnose. Der Nachweis der Enzymaktivität erfolgte durch Nachweis der Lycopinsynthese via HPLC. Die Klonierungsschritte sind im folgenden beschrieben:
- 25 **PCR 1.1:**
- Ca. 0,5 µg cDNA von *Blakeslea trispora*, 0,25 µM MAT350 5'-ACTTTATTGGATCCTTAAATGCGAATATCGTTGCTGC-3' (SEQ ID 58),
- 0,25 µM MAT244 5'-GTTCCAATTGGCCACATGAAGAGTAAGACAGGAAACAG-3' (SEQ ID
- 30 59), 100 µM dNTP, 10 µl Pfu-Polymerase-Puffer (10x), 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und H₂O ad 100µL.

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 40 °C 30s, 4. 72 °C 1 min 30 s, 5. 95 °C 30 s, 6. 50 °C 30 s, 7. 72 °C 1 min 30 s, 8. 95 °C 30 s, 9. 72 °C 10min
 Zyklen: (1-2.) 1x, (3-5.) 5x, (6-8.) 25x, (9.) 1x

5

PCR1.2:

Ca. 0,5 µg cDNA von *Blakeslea trispora*, 0,25 µM MAT243 5'-CCTGTCTTACTCTTCATGTGGCCAATTGGAACCAACAC-3' (SEQ ID 60), 0,25 µM MAT353 5'-CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3' (SEQ ID 61), 100 µM dNTP, 10 µl Pfu-Polymerase-Puffer (10x), 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und H₂O ad 100 µL.

10

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 40 °C 30s, 4. 72 °C 1 min 30 s, 5. 95 °C 30 s, 6. 50 °C 30 s, 7. 72 °C 1 min 30 s, 8. 95 °C 30s, 9. 72 °C 10min
 Zyklen: (1 -2.) 1x, (3-5.) 5x, (6-8.) 25x, (9.) 1x

15

Reinigung der PCR-Fragmente aus PCR 1.1, 1.2

Dazu wurde PCR 2 zur Herstellung der codierenden Sequenz von *carB* aus *Blakeslea trispora* für die Klonierung in pJOE2702 durchgeführt:

20

Ca. 50 ng Produkt aus PCR 1.1 und ca. 50 ng Produkt aus PCR1.2 mit 0,25 µM MAT350 (5'-ACTTTATTGGATCCTTAAATGCGAATATCGTTGCTGC-3' SEQ ID NO 58), 0,25 µM MAT353 (5'-CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3' SEQ ID NO 61), 100 µM dNTP, 10 µL Pfu-Polymerase-Puffer (10x), 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und H₂O ad 100 µL.

25

Temperaturprofil:

1. 95°C 10 min, 2. 85 °C 5 min, 3. 59 °C 30 s, 4. 72 °C 2 min, 5. 95 °C 30 s, 6. 72°C 10 min
 Zyklen: (1-2.) 1x, (3-5.) 22x, (6.) 1x

30

Anschließend erfolgte eine Reinigung des erhaltenen Fragmentes (~ 1,7 kbp), eine Ligation in Vektor pPCR-Script-Amp, eine Klonierung in *Escherichia coli* XL1-Blue, Sequenzierung der Insertion, Spaltung mit NdeI und BamHI sowie eine Ligation in pJOE2702. Das erhaltene Plasmid wurde als pBT4 bezeichnet.

Charakterisierung und Nachweis der Enzymaktivität von CarB (Phytoendesaturase)

Das von carB abgeleitete Genprodukt wurde als CarB bezeichnet. CarB weist auf Grundlage der Peptidsequenzanalyse folgende Eigenschaften auf:

Länge:	582 Aminoacylreste
Molekulare Masse:	66470
Isoelektrische Punkt:	6,7
Katalytische Aktivität:	Phytoendesaturase
Edukt:	Phytoen
Produkt:	Lycopin
EC-Nummer:	EC 1.14.99-

Der Nachweis der Enzymaktivität erfolgte in vivo. Wenn das Plasmid (pCAR-AE) in *Escherichia coli* XL1-Blue übertragen wird, entsteht der Stamm *Escherichia coli* XL1-Blue (pCAR-AE). Dieser Stamm synthetisiert Phytoen. Wenn zusätzlich das Plasmid pBT4 in *Escherichia coli* XL1-Blue übertragen wird, entsteht der Stamm *Escherichia coli* XL1-Blue (pCAR-AE)(pBT4). Da ausgehend von carB eine enzymatisch aktive Phytoendesaturase gebildet wird, produziert dieser Stamm Lycopin.

Die Plasmide pCAR-AE und pBT4 wurden daher in *Escherichia coli* übertragen. Nach Wachstum in Flüssigkultur wurden die Carotinoide aus den Zellen extrahiert und charakterisiert (vgl. oben).

Durch HPLC Analyse wurde nachgewiesen, daß der Stamm *Escherichia coli* XL1-Blue (pCAR-AE) Phytoen und der Stamm *Escherichia coli* XL1-Blue (pCAR-AE)(pBT4) Lycopin produziert. CarB weist folglich die Enzymaktivität einer Phytoendesaturase auf.

5

Herstellung gentechnisch veränderter Stämme von *Blakeslea trispora* zur Herstellung von Phytoen

Nachfolgend werden beispielhaft die Herstellung von gentechnisch veränderten Organismen zur Herstellung von Phytoen beschrieben.

10

Vector pBinAHygΔcarB zur Erzeugung von carB⁻ -Mutanten von *Blakeslea trispora*

Für die Deletion von carB in *Blakeslea trispora* wurde der Vektor pBinAHygΔcarB (SEQ. ID. NO:62, Fig. 22) konstruiert. Der Vorläufer von
15 pBinAHygΔcarB ist pBinAHyg (SEQ. ID. NO:3, Fig. 2). pBinAHyg wurde folgendermaßen konstruiert:

Aus dem Plasmid pANsCos1 (SEQ. ID. NO:4, Fig. 1, Osiewacz, 1994, Curr. Genet. 26:87-90) wurde die gpdA-hph Kasette als BglII/HindIII
20 Fragment isoliert und in das BamHI/HindIII geöffnete binäre Plasmid pBin19 (Bevan, 1984, Nucleic Acids Res. 12:8711-8721) ligiert. Der so erhaltene Vektor wurde als pBinAHyg bezeichnet und enthält das *E. coli* Hygromycin-Resistenzgen (hph) unter Kontrolle des gpd Promotors und des trpC Terminators aus *Aspergillus nidulans* sowie die entsprechenden Bordersequenzen, die für den DNA-Transfer von *Agrobacterium*
25 notwendig sind.

Die Amplifikation der codierenden Sequenz von carB mit den Primern MAT350 und MAT353 mittels PCR wurde mit den folgenden Parametern durchgeführt:

30 50 ng pBT4 mit 0,25 µM MAT350 (5'-ACTTTATTGGATCCTTAAAT-GCGAATATCGTTGCTGC-3'; SEQ ID NO 58), 0,25 µM MAT353 (5'-

CTATTTTAATCATATGTCTGATCAAAGAAGCATATTG-3'; SEQ ID NO 61), 100 µM dNTP, 10 µL Pfu-Polymerase-Puffer, 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 µL H₂O

Temperaturprofil:

- 5 1. 95 °C 10 min, 2. 85 °C 5 min, 3. 58 °C 30s, 4. 72 °C 2 min, 5. 95 °C 30s, 6. 72 °C 10 min.

Zyklen: (1.-2.) 1x, (3-5.) 30x, (6.) 1x

- Anschließend erfolgte eine Reinigung des erhaltenen Fragmentes (~ 1,7 kbp), eine Spaltung mit HindIII, eine weitere Reinigung des 364-bp-HindIII-Fragments-carB, gefolgt von einer Spaltung von pBinAHyg mit HindIII, eine Ligation von 364-bp-HindIII-Fragments-carB in pBinAHyg, eine Transformation des Vektors in *Escherichia coli* und eine Isolierung des Konstruktes und Bezeichnung als pBinAHygΔcarB wie oben beschrieben.
- 15 Alternativ erfolgte eine partielle Spaltung mit HindIII und die Klonierung eines größeren HindIII-Fragmentes aus carB in pBinAHyg zur Herstellung von pBinAHygΔcarB.

Erzeugung von carB⁻-Mutanten von *Blakeslea trispora*

- 20 Zunächst wurde das Plasmid pBinAHygΔcarB in den Agrobakterienstamm LBA 4404 übertragen, z. B. durch Elektroporation (vgl. oben). Anschließend wurde das Plasmid von *Agrobacterium tumefaciens* LBA 4404 in *Blakeslea trispora* ATCC 14272 und in *Blakeslea trispora* ATCC 14271 übertragen (vgl. oben). Der erfolgreiche Nachweis des
- 25 Gentransfers in *Blakeslea trispora* erfolgte über Polymerase-Kettenreaktion nach folgendem Protokoll:
- Ca. 0,5 µg DNA aus *Blakeslea trispora* ATCC 14272 carB⁻ bzw. ATCC 14271 carB⁻ wurden mit 0,25 µM Primer hph forward (5'-CGATGTAGGAGGGCGTGGATA-3'; SEQ ID NO 5), 0,25 µM Primer hph
- 30 reverse (5'-GCTTCTGCGGGCGATTGTGT-3'; SEQ ID NO 6), 100 µM

dNTP, 10 µL Herculanse-Polymerase-Puffer, 2,5 U Herculanse-DNA-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 µl H₂O umgesetzt.

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 58 °C 1 min, 4. 72 °C 1 min, 5. 94 °C 1 min, 6. 72 °C 10 min.

Zyklen: (1.-2.) 1x, (3.-5.) 30x, (6.) 1x

Als Negativkontrolle wurde eine Amplifikation des Kanamycinresistenzgens aus *Agrobacterium* versucht. Dazu wurden folgende PCR-Bedingungen verwendet:

Ca. 0,5 µg DNA aus *Blakeslea trispora* ATCC 14272 carB⁻ bzw. ATCC 14271 carB⁻ wurden mit 0,25 µM Primer nptIII forward (5'-TGAGAATATCACCGGAATTG-3'; SEQ ID NO 7), 0,25 µM Primer nptIII reverse (AGCTCGACATACTGTTCTTCC-3'; SEQ ID NO 8), 100 µM

dNTP, 10 µL Herculanse-Polymerase-Puffer, 2,5 U Herculanse-DNA-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 µL H₂O umgesetzt.

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 58 °C 1 min, 4. 72 °C 1 min, 5. 94 °C 1 min, 6. 72 °C 10 min-

Zyklen: (1.-2.) 1x, (3.-5.) 30x, (6.) 1x

Produktion von Carotinoiden und Carotinoidvorstufen mit *Blakeslea trispora*

Zur Produktion der Carotinoide Zeaxanthin, Canthaxanthin, Astaxanthin und Phytoen wurden die entsprechenden gentechnisch veränderten *Blakeslea trispora* (+) und (-) Stämme fermentiert, das produzierte Carotinoid mittels HPLC Analyse nachgewiesen und isoliert.

Das Flüssigmedium zur Produktion von Carotinoiden enthielt pro Liter: 19 g Maismehl, 44 g Sojamehl, 0,55 g KH_2PO_4 , 0,002 g Thiaminhydrochlorid, 10 % Sonnenblumenöl. Der pH wurde mit KOH auf 7,5 eingestellt.

- 5 Zur Herstellung der Carotinoiden wurden Schüttelkolben mit Sporensuspensionen von (+) und (-) Stämmen der GVO von *Blakeslea trispora* beimpft. Die Schüttelkolben wurden bei 26 °C mit 250 rpm für 7 Tage inkubiert. Alternativ wurde zu Mischungen der Stämme nach 4 Tagen Trisporensäuren zugegeben und weitere 3 Tage inkubiert. Die
10 Endkonzentration der Trisporensäuren betrug 300 - 400 µg/ml.

Extraktion und Analytik

Extraktion:

1. Entnahme von 10 ml Kultursuspension
- 15 2. Zentrifugation, 10 min, 5.000 x g
3. Verwerfen des Überstandes
4. Resuspendierung des Pellets in 1 ml Tetrahydrofuran (THF) durch Vortexen
5. Zentrifugation, 5 min, 5.000 x g
- 20 6. Abnahme der THF-Phase
7. Wiederholung der Schritte 4.-6. (2 x)
8. Vereinigung der THF-Phasen
9. Zentrifugation der vereinigten THF-Phasen 5 min bei 20.000 x g, um Reste der wäßrigen Phase abzutrennen

25

Analytik

Messung von Phytoen mittels HPLC

- Säule: ZORBAX Eclipse XDB-C8, 5 µm, 150*4,6 mm
Temperatur: 40 °C
30 Flußrate: 0,5 ml/min
Injektionsvolumen: 10 µl

Detektion: UV 220 nm
 Stoppzeit: 12 min
 Nachlaufzeit: 0 min
 Maximaldruck: 350 bar
 5 Eluent A: 50 mM NaH₂PO₄, pH 2,5 mit Perchlorsäure
 Eluent B: Acetonitril
 Gradient:

	Zeit [min]	A [%]	B [%]	Fluß [ml/min]
	0	50	50	0,5
10	12	50	50	0,5

Als Matrix wurden Extrakte der Fermentationsbrühen verwendet. Vor der HPLC wurde jede Probe durch ein 0,22 µm Filter filtriert. Die Proben wurden kühl gehalten und vor Licht geschützt. Zur Kalibrierung wurden
 15 jeweils 50 - 1000 mg/l eingewogen und in THF gelöst. Als Standard wurde Phyton verwendet, welches unter den gegebenen Bedingungen eine Retentionszeit von 7,7 min. aufweist.

Messung von Lycopin, β-Carotin, Echinenon, Canthaxanthin, Cryptoxanthin, Zeaxanthin und Astaxanthin mittels HPLC

20 Säule: Nucleosil 100-7 C18, 250*4,0 mm (Macherey & Nagel)
 Temperatur: 25 °C
 Flußrate: 1,3 ml/min
 Injektionsvolumen: 10 µl
 25 Detektion: 450 nm
 Stoppzeit: 15min
 Nachlaufzeit: 2 min
 Maximaldruck: 250 bar
 Eluent A: 10% Aceton, 90% H₂O
 30 Eluent B: Aceton
 Gradient:

	Zeit [min]	A [%]	B [%]	Fluß [ml/min]
	0	30	70	1,3
	10	5	95	1,3
	12	5	95	1,3
5	13	30	70	1,3

Als Matrix wurden Extrakte der Fermentationsbrühen verwendet. Vor der HPLC wurde jede Probe durch ein 0,22 µm Filter filtriert. Die Proben wurden kühl gehalten und vor Licht geschützt. Zur Kalibrierung wurden jeweils 10 mg eingewogen und in 100 ml THF gelöst. Als Standard wurden folgende Carotinoide mit folgenden Retentionszeiten eingesetzt β-Carotin (12,5 min), Lycopin (11,7 min), Echinenon (10,9 min), Cryptoxanthin (10,5 min), Canthaxanthin (8,7 min), Zeaxanthin (7,6 min) und Astaxanthin (6,4 min) [s. Fig 23].

15

Produktion von Zeaxanthin mit gentechnisch veränderten Stämmen von *Blakeslea trispora*

Nachfolgend wird beispielhaft die Herstellung von Zeaxanthin mit gentechnisch veränderten Organismen (GVO) von *Blakeslea trispora* beschrieben.

20

Durch Agrobakterium-vermittelte Transformation wurde der Vektor pBinAHygBTpTEF1-HPcrtZ in *Blakeslea trispora* übertragen (s.o.). Ein Hygromycin-resistenter Klon wurde isoliert und auf eine Kartoffel-Glucose-Agarplatte (Merck KGaA, Darmstadt) übertragen.

25

Nach drei Tagen Inkubation bei 26°C wurde ausgehend von dieser Platte eine Sporensuspension hergestellt. Ein 250-ml-Erlenmeyerkolben ohne Schikanen mit 50 ml Growth-Medium (Maismehl 47 g/l, Sojamehl 23 g/l, KH₂PO₄ 0,5 g/l, Thiamin-HCl 2,0 mg/l, pH mit NaOH vor der Sterilisation auf 6,2–6,7 eingestellt) wurde mit 1x10⁸ Sporen beimpft.

30

Diese Vorkultur inkubierte 48 Stunden bei 26 °C und 250 upm. Für die Hauptkultur wurde ein 250-ml-Erlenmeyerkolben ohne Schikane

enthaltend 40 ml Produktionsmedium mit 4 ml der Vorkultur beimpft und 8 Tage bei 26 °C und 150 upm inkubiert. Das Produktionsmedium enthielt Glucose 50 g/l, Casein Acid Hydrolisat 2 g/l, Hefeextrakt 1 g/l, L-Asparagin 2 g/l, KH_2PO_4 1,5 g/l, $\text{MgSO}_4 \times 7 \text{ H}_2\text{O}$ 0,5 g/l, Thiamin-HCl 5 mg/l, Span20 10 g/l, Tween 80 1 g/l, Linolsäure 20 g/l, Maisquellwasser 80 g/l. Nach 72 Stunden erfolgte die Zugabe von Kerosin in einer Endkonzentration von 40 g/l Kerosin.

Nach der Ernte der Kulturen werden die verbliebenen ungefähr 35 ml Kultur mit Wasser auf 40 ml aufgefüllt. Anschließend werden die Zellen im Hochdruckhomogenisator, Typ Micron Lab 40, Fa. APV Gaulin, 3 x bei 1500 bar aufgeschlossen.

Die Suspension mit den aufgeschlossenen Zellen wurde mit 35 ml THF versetzt und 60 min bei RT im Dunkeln bei 250 upm geschüttelt. Danach wurden 2 g NaCl zugegeben und das Gemisch nochmals geschüttelt. Der Extraktionsansatz wurde dann 10 min bei 5000 x g zentrifugiert. Die gefärbte THF-Phase wurde abgenommen, die Zellmasse war vollständig entfärbt.

Die THF-Phase wurde am Rotationsverdampfer bei 30 mbar und 30 °C auf 1 ml eingengt und danach nochmals in 1 ml THF aufgenommen. Nach Zentrifugation 5 min bei 20 000 x g wurde ein Aliquot der oberen Phase entnommen und durch HPLC analysiert (Fig. 24, Fig. 23).

Patentansprüche

1. Verfahren zur Herstellung eines gentechnisch veränderten Organismus der Gattung *Blakeslea* umfassend
 - 5 (i) Transformation mindestens einer der Zellen,
 - (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder mehreren genetischen Merkmalen alle gleichartig verändert sind und diese gentechnische Veränderung zur Ausprägung bringen, und
 - 10 (iii) Selektion und Anzucht der gentechnisch veränderten Zelle oder Zellen.
2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** es sich um Zellen von Pilzen der Art *Blakeslea trispora* handelt.
3. Verfahren nach Anspruch 1 oder 2, **dadurch gekennzeichnet, dass** in
15 der Transformation (i) ein Vector oder freie Nukleinsäuren verwendet werden.
4. Verfahren nach Anspruch 3, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector in das Genom mindestens einer der Zellen integriert wird.
- 20 5. Verfahren nach Anspruch 4, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector einen Promotor und/oder einen Terminator enthält.
6. Verfahren nach einem der vorhergehenden Ansprüche 3 bis 5, **dadurch gekennzeichnet, dass** in der Transformation (i) ein Vector
25 enthaltend den *gpd*, *pcarB*, *pcarRA* und/oder *ptef1* Promotor und/oder den *trpC* Terminator eingesetzt wird.

7. Verfahren nach einem der vorhergehenden Ansprüche 3 bis 6, **dadurch gekennzeichnet, dass** in der Transformation (i) ein Vector enthaltend ein Resistenzgen eingesetzt wird.
- 5 8. Verfahren nach Anspruch 7, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector ein Hygromycin-Resistenzgen (hph), insbesondere aus *E. coli* enthält.
9. Verfahren nach einem der vorhergehenden Ansprüche 5 - 8, **dadurch gekennzeichnet, dass** der *gpd* Promotor die Sequenz SEQ ID NO: 1
- 10 aufweist.
10. Verfahren nach einem der vorhergehenden Ansprüche 5 - 8, **dadurch gekennzeichnet, dass** der *trpC* Terminator die Sequenz SEQ ID NO: 2 aufweist.
11. Verfahren nach einem der vorhergehenden Ansprüche 5 - 8, **dadurch**
- 15 **gekennzeichnet, dass** der *tef1* Promotor die Sequenz SEQ ID NO: 35 aufweist.
12. Verfahren nach einem der Ansprüche 6 bis 11, **dadurch gekennzeichnet, dass** der *gpd* Promotor und der *trpC* Terminator aus *Aspergillus nidulans* stammen.
- 20 13. Verfahren nach einem Ansprüche 3 bis 12, **dadurch gekennzeichnet, dass** der Vector die SEQ ID NO: 3 umfasst.
14. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die Transformation (i) mittels Agrobakterien, Konjugation, Chemikalien, Elektroporation, Beschuss mit DNA-beladenen Partikeln, Protoplasten oder Mikroinjektion durchgeführt
- 25 wird.

15. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** in der Homokaryontisierung (ii) ein mutagenes Agens eingesetzt wird.
16. Verfahren nach Anspruch 15, **dadurch gekennzeichnet, dass** als mutagenes Agens N-Methyl-N'-nitro-nitrosoguanidin (MNNG), UV-Strahlung oder Röntgenstrahlung eingesetzt wird.
17. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die Selektion durch Markierung und/oder Auswahl der einkernigen Zellen erfolgt.
18. Verfahren nach einem der vorhergehenden Ansprüche 1 - 17, **dadurch gekennzeichnet, dass** in der Selektion 5-Carbon-5-deazariboflavin (darf) und Hygromycin (hyg) oder 5-Fluororotat (FOA) und Uracil und Hygromycin eingesetzt werden.
19. Verfahren nach einem der Ansprüche 3 bis 18, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Carotinoiden oder deren Vorstufen enthält.
20. Verfahren nach einem der Ansprüche 3 bis 19, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Carotinen oder Xanthophyllen enthält.
21. Verfahren nach einem der Ansprüche 3 bis 20, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Adonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, β -Carotin, α -Carotin, Lutein, Bixin, Phytofluen oder Phytoen enthält.

22. Verfahren nach einem der Ansprüche 3 bis 21, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector derart gestaltet ist, dass die im Vector enthaltene genetische Information in das Genom von *Blakeslea trispora* eingeführt wird
- 5 23. Verfahren nach einem der Ansprüche 3 bis 22, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen enthält, die nach Expression eine Ketolase- und/oder Hydroxylase-Aktivität entfalten.
24. Verfahren nach Anspruch 23, **dadurch gekennzeichnet, dass** der in
10 der Transformation (i) eingesetzte Vector die SEQ ID NO: 70 oder 71 oder 76 und/oder 72 umfasst.
25. Verfahren nach Anspruch 23 oder 24, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector eine Sequenz aus der Gruppe bestehend aus den SEQ ID NO: 37 – 51 aufweist.
- 15 26. Verfahren nach einem der Ansprüche 3 bis 21, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector derart gestaltet ist, dass die im Vector enthaltene genetische Information in der Zelle ausgeschaltet wird.
27. Verfahren nach einem der Ansprüche 3 bis 21 oder 25, **dadurch**
20 **gekennzeichnet, dass** durch die Transformation (i) das Gen der Phytoendesaturase ausgeschaltet wird.
28. Verfahren nach Anspruch 27, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector die **SEQ ID NO: 69** umfasst.
29. Verfahren nach Anspruch 27 oder 28, **dadurch gekennzeichnet, dass**
25 der in der Transformation (i) eingesetzte Vector die Sequenz SEQ ID NO: 62 aufweist.

30. Verfahren nach einem der Ansprüche 3 bis 21, **dadurch gekennzeichnet, dass** durch die Transformation das Gen der Lycopincyclase ausgeschaltet wird.
- 5 31. Genetisch veränderte mehrkernige Zellen der zur Gattung *Blakeslea* gehörenden Pilze, insbesondere *Blakeslea trispora* erhältlich nach einem der vorhergehenden Ansprüche.
32. Verwendung der Zellen nach Anspruch 30 oder eines aus ihnen gebildeten Mycels zur Herstellung von Carotinoiden oder deren Vorstufen.
- 10 33. Verwendung nach Anspruch 30 oder 31 zur Herstellung von Carotinene oder Xanthophyllen.
34. Verwendung nach einem der Ansprüche 30 bis 32 zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Anthonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-
15 Hydroxyechinenon, Lycopin, β -Carotin, α -Carotin, Lutein, Bixin, Phytofluene oder Phytoen.
35. Promotor mit der Sequenz SEQ ID NO: 1 oder 35 zur Verwendung in dem Verfahren nach einem der Ansprüche 1 - 29.
- 20 36. Terminator mit der Sequenz SEQ ID NO: 2 zur Verwendung in dem Verfahren nach einem der Ansprüche 1 - 29.
37. Vector umfassend die SEQ ID NO: 3 zur Verwendung in dem Verfahren nach einem der Ansprüche 1 - 29.
38. Vector nach Anspruch 36 zur Verwendung in dem Verfahren nach einem der Ansprüche 1 - 29 umfassend die SEQ ID NO: 69 und/oder
25 die SEQ ID NO: 70 oder 71 und/oder 72 oder 76.

Fig. 1: Vektor pANsCos1

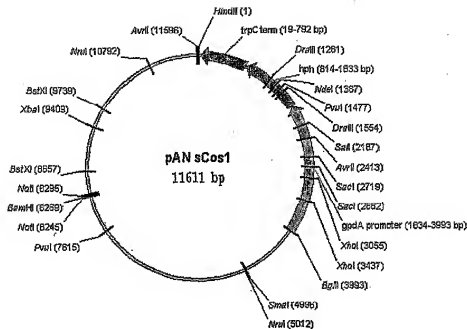


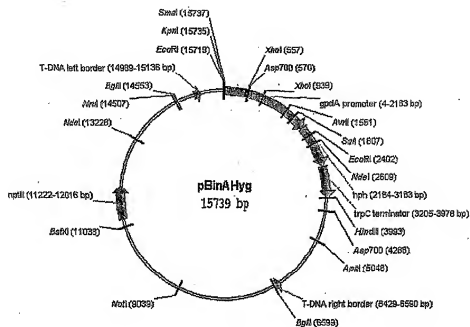
Fig. 2: Vektor pBinAHyg

Fig. 3: Gels des Ergebnis einer PCR

Spur:

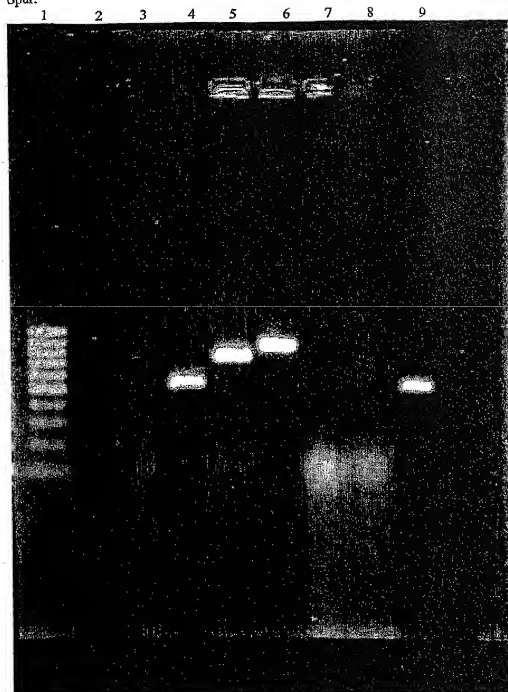


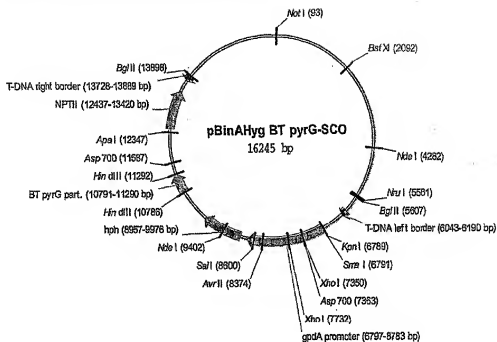
Fig. 4: Plasmid pBinAHygBTpyrG-SCO

Fig. 5: Plasmid pBinAHygBTpTEF1-HPcrtZ

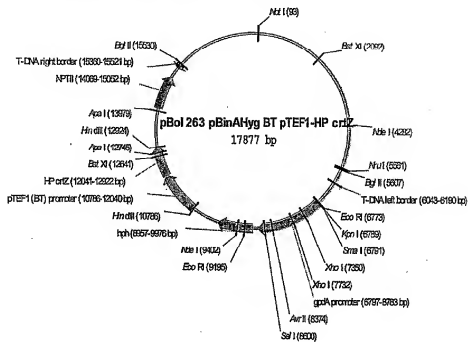


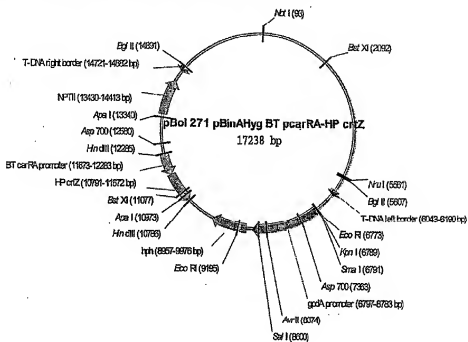
Fig. 6: Plasmid pBinAHyg-BTpcarRA-HPcrtZ

Fig. 7: Plasmid pBinAHygBTpcarB-HPcrIZ

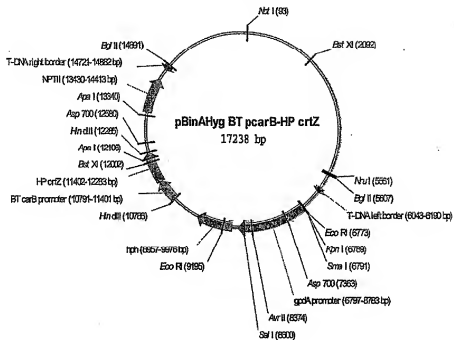


Fig. 8: Plasmid p-carRA-HPcrtZ-TAG-3'carA-IR

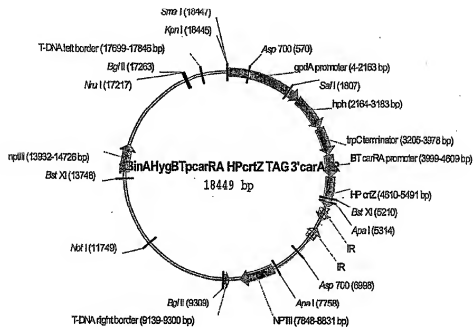


Fig. 10: Plasmid pBinAHygBTpTEF1-EUcrtZ

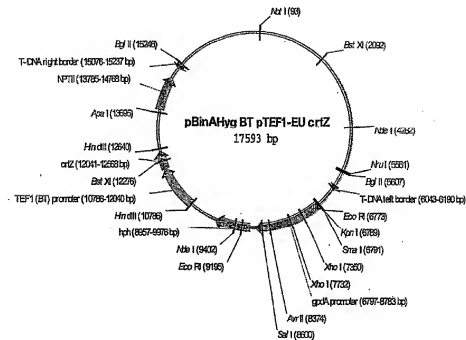


Fig. 11: Plasmid pBinAHygBTpcarRA-EUcrtZ

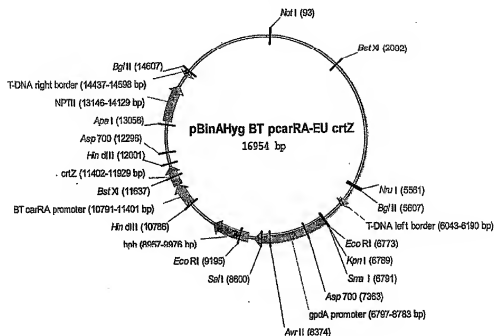


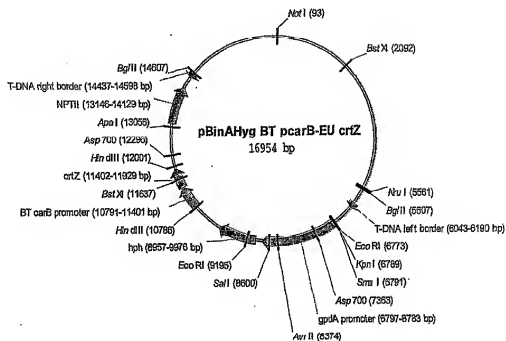
Fig. 12: Plasmid pBinAHyg BT pcarB-EU crtZ

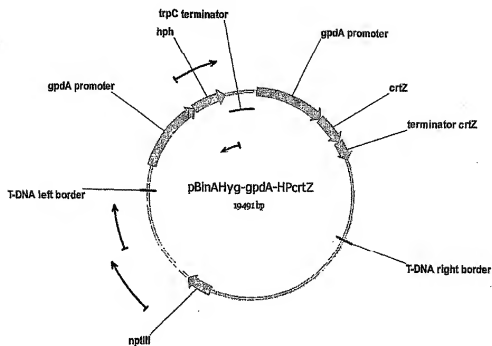
Fig. 13: Plasmid p-BinAHyg-gpdA-HPcrtZ

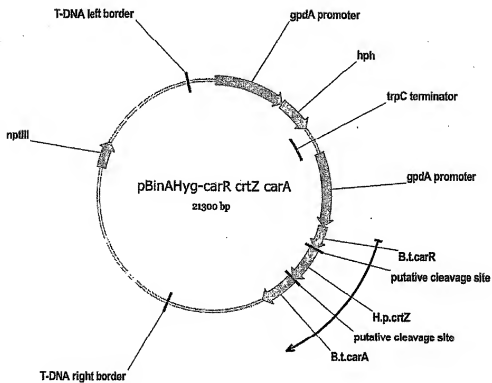
Fig. 14: Plasmid pBinAHyg-carRcrtZcarA

Fig. 15: Plasmid pBinAHyg-BTpTEF1-NPcrtW

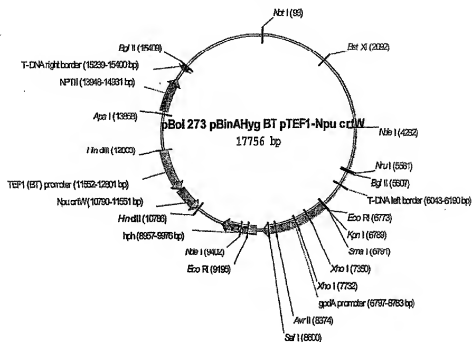


Fig. 16: Plasmid pBinAHyg_BTpcarRA_NPcrtW

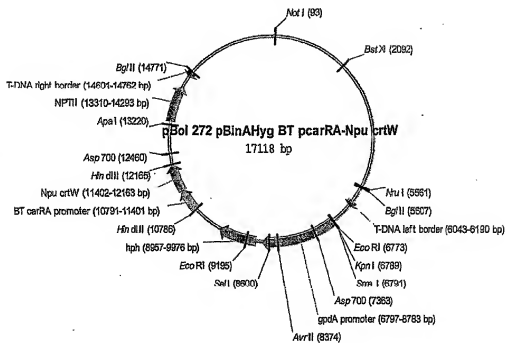


Fig. 17: Plasmid pBinAHyg-BTpcarB-NPcrtW

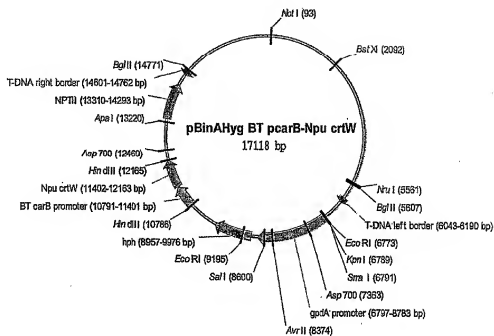


Fig. 18: Plasmid pBinAHygBTpcarRA-HPcrfZ-BTpcarRA-NpuclrtW

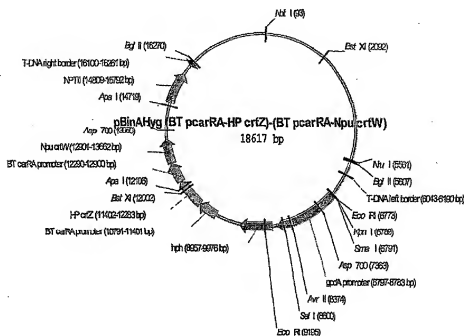


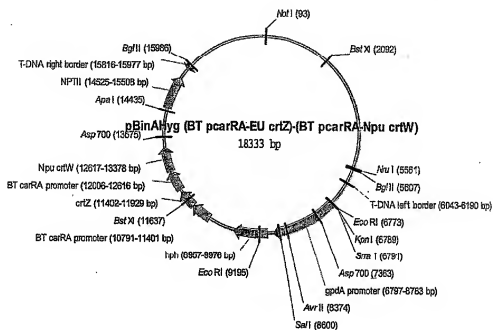
Fig. 19: Plasmid pBinAHygBTpcarRA-EUcrtZ-BTpcarRA-Npu crtW

Fig. 20: carB

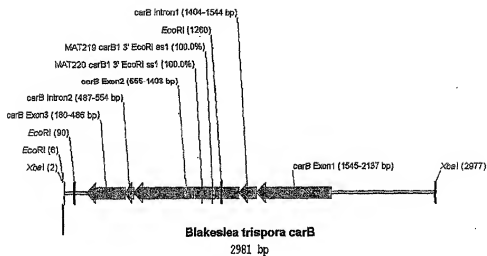


Fig. 21: CDS von carB



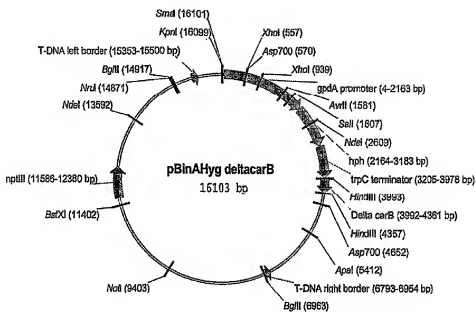
Fig. 22: Vektor pBinAHyg Δ carB

Fig. 23: HPLC Standard

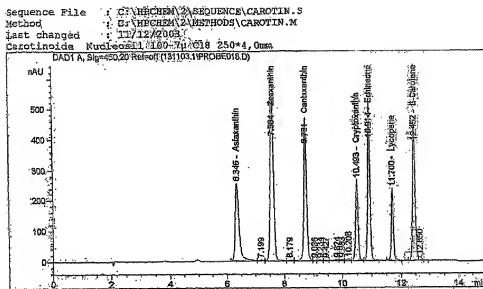
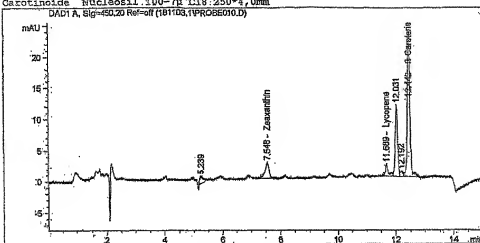


Fig. 24: HPLC

Sequence File : C:\HPCHEM\2\SEQUENCE\CAROTIN.S
Method : C:\HPCHEM\2\METHODS\CAROTIN.M
Last changed : 11/18/2003
Carotinoide Nucleosil 100-7, C18, 250*4, 0mm
DAD1 A, Sig=450.20 Ref=off (101103, VPROB010.D)



SEQUENCE LISTING

<110> BASF AG

<120> Verfahren zur gentechnischen Veränderung von Organismen der Gattung *Blakeslea*, entsprechende Organismen und deren Verwendung

<130> ?

<160> 80

<170> PatentIn version 3.2

<210> 1

<211> 2160

<212> DNA

<213> Artificial

<220>

<223> Promotor

<400> 1

```

ctttcgacac tgaataagct agagcctgct oagcttgaa gggcgagga gctcgtact      60
gtcacaaacta ccaacatgga gtacgataag ggccagttcc gccagctcat taagagccag      120
ttcatgggog ttggcatgat ggccgtcatg catctgtact tcaagtagac caacgctctt      180
ctgatccagt cgatcatccg ctgaaggcgc tttcgaatct ggttaagatc caagctcttcg      240
ggaagccagc gactggtagc ctccagcgtc cctttaaggc tgccaacagc tttctcagcc      300
agggccagcc caagaccgac aaggcctccc tccagaacgc cgagaagaac tggaggggtg      360
gtgtcaagga ggagtaagct ccttattgaa gtccgaggac ggagcggtag caagaggata      420
ttcttcgact ctglalata gataagatga tgaggaattg gggtagcat agcttcattt      480
ggatttgctt tccaggtgta gactctagct tggagcatag agggctcttt ggctttcaat      540
attctcaagt atctcgagtt tgaacttatt ccctgtgaac cttttattca ccaatgagca      600

```

ttggaatgaa catgaatctg aggactgcaa togccatgag gttttcgaaa tacatccgga 660
 tgtcgaaggc ttggggcacc tgcgttggtt gaatttagaa cgtggcacta ttgatcatcc 720
 gatagctctg caaagggcgt tgcacaatgc aagtcacacg ttgctagcag ttccaggctg 780
 aatgttatga tgcagattgt attaaatcag gagatatagc atgatctcta gttagctcac 840
 cccaaagtc agacggcgta accaaaagtc acacaacana agctgtaagg atttcacac 900
 ggctacggaa gacggagaag ccaccttcag tggactcgag taccatttaa ttctatttgt 960
 gtttgatcga gacctaatcc agcccctaca acgacctca aagtcgtata gctaccagtg 1020
 aggaagtgga ctcaaatoga cttcagcaac atctcctgga taaactttaa gcctaaacta 1080
 tacagaataa gatagggtga gagcttatac cgagctcca aatctgtcca gatcatggtt 1140
 gaccggtgcc tggatcttcc tatagaatca tcttatttcg ttgacctagc tgattctgga 1200
 gtgaccaga gggctcatgac ttgagcctaa aatccgcgcg ctccaccatt tctagaaaaa 1260
 tgtgacgaac tcylgagctc tgtacagtga ccggtgactc tttctggcat cgggagagac 1320
 ggacggacgc agagagaagg gctgagtaat aagccactgg ccagacagct ctggcggtcc 1380
 tgagggtcgag tggatgatta ttaatccggg accggccgcc cctccgcccc gaagtggaaa 1440
 ggctggtgtg cccctcgttg accaagaato tattgcatca toggagaata tggagcttca 1500
 tcgaatcacc ggcagtaagc gaaggagaat gtgaagccag ggggtgtatag ccgtcggcga 1560
 aatagcatgc cattaacctc ggtacagaag tccaattgct tcgatctggt taaaagattc 1620
 acgagatagt accttctcgg aagtaggtag agcgagtacc cggcggttaa gctccctaatt 1680
 tggccatcc ggcattctga gggcgtccaa atalülycc tctcctgctt tgcgggtgt 1740
 atgaaacggg aaaggccgct caggagctgg ccagcggcgc agaccgggaa cacaagctgg 1800
 cagtcgaacc atccggtgct ctgcactoga cctgctgagg tccctcagtc cctggttaggc 1860

agcttttgccc cgtctgtccg cccggtgtgt cggcgggggtt gacaaggctg ttgcgtcagt 1920

ccaacatttg ttgccatatt ttcctgctct ccccaccagc tgctcttttc tttctctttt 1980

cttttcccat cttcagtata ttcctcttcc catccaagaa cttttatttc ccttaagtaa 2040

gtacttttgct acatccatac tccatcttcc ccatccctta ttcttttgaa cttttcagtt 2100

cgcgctttcc caottcctag cagcttgact aacagctacc ccgcttgagc agacatcac 2160

<210> 2

<211> 774

<212> DNA

<213> Artificial

<220>

<223> Terminator

<220>

<221> misc_feature

<222> (267)..(267)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (475)..(475)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (566)..(566)

<223> n is a, c, g, or t

<400> 2

cgatccactt aacgttactg aaatcatcaa acagcttyac gaatctggat ataagctgt 60

tggtgtcgat gtcagctcgg gagttgagac aaatggtggt caggatctcg ataagatag 120

ttcatttgtc caagcagcaa agagtgcctt ctagtgattt aatagctcca tgtaacaag 180

.aataaaacgc gttttcgggt ttacctcttc cagatacagc tcctctgcaa tgcattaatg 240
cattgactgc aacctagtaa cgccttncag gctcggcga agagaagaat agcttagcag 300
agctattttc attttcggga gacgagatca agcagatcaa cggtcgtcaa gagacctacg 360
agactgagga atccgctctt ggcctccagc gactatataat ttgtctctaa ttgtactttg 420
acatgctcct cttctttact ctgatagctt gactatgaaa attccgtcac cagcncctcg 480
gttcgcgaag ataattgcat gtttcttctt tgaactctca agcctacagg acacacattc 540
atcgtaggta taaacctcga aatcanttcc tactaagatg gtatacaata gtaaccatgc 600
atgggttgct agtgaatgct ccgtaacacc caatacgcg gccgaaactt ttttacaact 660
ctctatgag tcgtttaacc agaatgcaca ggtacacttg tttagaggta atccttcttt 720
ctagctagaa gtctcgtgt actgtgtaag cgccactcc acatctccac tcga 774

<210> 3

<211> 15739

<212> DNA

<213> Artificial

<220>

<223> Vector

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 3

gatcttttcca cactgaaata cgtcgagcct gtcctcgcttg gaagcggcga ggagcctcgt	60
cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc	120
cagttcatgg gcgttggcat gatggcngtc atgcctctgt acttcaagta caccaacgct	180
cttctgatcc agtcgataat ccgctgaagg cgttttcgaa totggttaag atccacgtct	240
tccgggaagcc agcgcactggt gacctccagc gtccctttaa ggctgccaac agctttctca	300
gccaggggcca gcccaagacc gacaaggcct cctccagaa cgcgcagaag aactggaggy	360
gtggtgtcaa ggaggagtaa gtccttatt gaagtcggag gacggagcgg tgtcaagagg	420
atattcttgg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca	480
tttgattttg ctttccaggg tgagactcta gtttggagca tagagggtcc tttggcttcc	540
aatattctca agtatctoga gtttgaactt attcctgtg aaccttttat tccaactga	600
gcatttgaat gaacatgaat ctgaggactg caatcgccat gaggttttcc aaatacatcc	660
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca	720
tcgatagctc ctgcaagggg cgttgcacaa tgcaagtcga acgttgctag cagttccagg	780
tggaaatgta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct	840
caccacaaaa gtcagacggc gtaacaaaa gtcacacaac acaagctgta aggatttcgg	900
cacggctacg gaagacggag aagccacctt cagtggaetc gagtaccatt taattctatt	960
tgtgtttgat cgagacataa taagacccct acaacgacca tcaagtcgt atagctacca	1020
gtgaggaagt ggactcaaat cgaattcagc aacatctcct ggataaactt taagcctaaa	1080
ctatacagaa taagataggt ggagagctta tacogagctc ccaaatctgt ccagatcatg	1140

gttgaccggt gcttgatctt tcctatagaa tcctccttat tcgttgacct agctgattct 1200
 ggagtgacc agagggatcat gacttgagcc taaaatccgc cgcctccacc atttgtagaa 1260
 aatgtgagc aactcgtgag ctctgtacag tgacoggtga ctctttctgg catgcggaga 1320
 gcgggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380
 ctctggagtg cagtggtga ttattaatcc gggacggcc gccctcgcgc ccgaagtgg 1440
 aaaggctggg gtgcccctcg ttgaccaaga atctattgca tcacoggaga atatggagct 1500
 tcactgaac accggcagta agcgaaggag aatgtgaagc caggggtgta tagcgcgcg 1560
 cgaaatagca tgccattaac ctaggtagc aagccaatt gctccgac tcgttaaaga 1620
 ttcacgagat agtacctctt ccgaagtag tagagcgagt accggcgcg taagctcct 1680
 aattggccca tcggcatctt gtaggcgctc caaatatcgt gcctctcctg ctttgccgg 1740
 tgtatgaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800
 tggcagtcga cccatccggl gclctgcact cgacctgctg aggtccctca gtccctggta 1860
 ggcagctttg ccccgctctt ccgcocgggt gtgcggcggg gttgacaagg tcgttgcgct 1920
 agtccaacat ttgttgccat attttcctgc tctcccacc agctgcctct tctttttctc 1980
 tttcttttcc catcttcagt atattcatct tcccaccaa gaacctttat ttcccctaag 2040
 taagtacttt gctacatcca tactocatcc ttccatccc ttattccttt gaacctttca 2100
 gttcgagctt tcccactcca tcgcagcttg actaacagct acccgcgttg agcagacatc 2160
 accatgcctg aactcaccgc gacgtctgtc gagaagtctc tgatcgaaaa gttcgacagc 2220
 gtctcggacc tgatgcagct ctggaggggc gaagaatctc gtgctttcag ttctgatgta 2280
 ggaggcgctg gatatgtcct gcgggtaaat agctgcgcgc atggtttcta caaagatcgt 2340
 tatgtttatc ggcactttgc atggcgcg ctcccgattc cggagtgct tgacattggg 2400

gaattcagcg agagcctgac ctattgcac tcgccgcgtg cacagggtgt caggttgcaa 2460

gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg 2520

atcgctgcgg cagatcttag ccagacgagc ggggttcggcc cattcggacc gcaaggaato 2580

ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640

tggcaaaactg tgatggagca caccgtcagt gcgtccgtcg cgcaggctnt cgtatgagctg 2700

atgctttggg ccgaggactg ccccgaaagt cggcaacctg tgcaacggga tttcggctcc 2760

aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820

ttcggggatt cccaatacga ggtcgccaac atottcttct ggaggccctg gttggcttgt 2880

atggagcagc agacgcgcta ctctgagcgg aggcacccgg agcttcgagg atocccgcgg 2940

ctccgggcgt atatgctccg cattggtctt gaccaactct atcagagctt ggttgacggc 3000

aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgtccg atocggagcc 3060

gggactctcy ggcgtacaca aatgccegc agaagcgagg ccgtctggac cgatggctgt 3120

gtagaagtac tcgccgatag tggaaaccga cgcgccagca ctcgtccgag ggcaaaggaa 3180

tagagtagat gcgcaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240

tgacgaatct ggataaaga togttggtgt cgatgtcagc tcggaggttg agacaaatgg 3300

tgttcaggat ctgcataaga tacgttcatt tgtccaagca gcaaagagtg ccttctagt 3360

atttaatatg tccatgtcaa caagaataaa acgcgttttc gggtttacct ctccagata 3420

cagctcatct gcaatgcatt aatgcattga ctgaaccta gtaacgcctt ncaggtccg 3480

gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540

tcaacggtcg tcaagagacc tacgagactg aggaatccgc tcttggtcc acgcgactat 3600

atatttgtct ctaattgtac ttgacatgc tcctcttctt tactctgata gcttgactat 3660

gaaaattécg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tecttgaact 3720
 ctcaagccta caggacacac attcatcgta ggtataaacc tcgaatcan ttctactaa 3780
 gatggttatc aatagtaacc atgcatggtt gcctagtga tgctccgtaa caccaatac 3840
 gccggccgaa acctttttac aactctccta tgagtcgttt acccagaatg cacaggta 3900
 ctgttttaga ggtaatcctt ctttctactt agaagtcctc gtgtactgtg taagcgcca 3960
 ctccacatct cactcgcac tgcaggcatg caagcttgcc gtaatcatgg tcatagtgt 4020
 ttctgtgtg aaattgttat ccgctcaca ttccacaca catacgagcc ggaagcata 4080
 agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaatggog ttgcgtcac 4140
 tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaac ggccaacgcg 4200
 cggggagagg cgttttgctt attgggcaa agacaaaagg ggcacattca accgattgag 4260
 ggagggaagg taaatatga cggaaattat tcattaaagg tgaattatca ccgtcaccga 4320
 cttgagccat ttgggaatta gagccagca aatcaccagt agcaccatta ccattagcaa 4380
 ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta ggcacagaat 4440
 caagtttgcc tttagcgtca gactgtagcg cgttttcac gccattttcg gtcatagcc 4500
 ccttattagc gtttgccac ttttcataat caaaatcac ggaaccagag ccaccaccg 4560
 aaccgcctcc ctcagagccg ccacctcag aaccgccacc ctcagagcca ccacctcag 4620
 agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggagcg ccgatctagt 4680
 aacatagatg acaccgcg cgataattta tctagtttg ccgcctatat tttgtttct 4740
 atugcgtatt aatgtataa ttgogggact ctaatcataa aaacctctc cataataac 4800
 gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa ttatatgata 4860
 atcatcgcaa gaccggaac aggatccaat ctaagaaac tttattgcca aatgtttgaa 4920

cgatcgggga tcatacgggt ctgtggcggg aactccacga aaatatccga acgcagcaag 4980
 atatcgcggt gcatctcggt ctgtccctggg cagtcgcgcg cgacgcggtt gatgtggagc 5040
 cggggcccca tcataattgtc gctcaggatc gtggcggttg gcttgctggc cgttgctgtc 5100
 gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccgtaggc gaagaactcc 5160
 agcatgagat ccccgccgtg gaggatcacc cagccggcgt cccgaaaac gattccgaag 5220
 cccaaccttt catagaaggc ggcggtgaa tcgaaatctc gtgatggcag gttgggcgtc 5280
 gcttggtcgg tcatttgaa cccagagtc ccgctcagaa gaactcgtca agaaggcgat 5340
 agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg aagcggtcag 5400
 cccattcgcc gccaaagctc tcagcaatat caccggtagc caacgctatg tcctgatagc 5460
 ggtccgccac accagccgg ccacagtcga tgaatccaga aaagcggcca tttccacca 5520
 tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg tcgggcctgc 5580
 ggccttgag cctggcgaa agttoggctg gcgcgagccc ctgatgctct tcgtccagat 5640
 catcctgac gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg 5700
 cttggtggtc gaatggcgag gtagccggat caagcgtatg cagccgccg attgcatcag 5760
 ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc tgcgccgca 5820
 cttcgccaa tagcagccag tcccttccg cttcagtgac aacgtcgagc acagctgcgc 5880
 aaggaaacgc cgtcgtggcc agccacgata gccgcgctc ctcgtcctgc agttcattca 5940
 gggcacccga caggtcggtc ttgacaaaaa gaaccgggcg cccctcgct gacagccgga 6000
 acacggcggc atcagagcag ccgattgtct gttgtgccca gtcacagccg aatagcctct 6060
 ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg cgaacgato 6120
 cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt ggataccgag 6180

gggaatttat ggaacgtcag tggagcatat ttgacaagaa atatttgcta gctgatagtg 6240

acottaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt agctcattaa 6300

actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca gttccaaacg 6360

taaaaoggct tgtcccgctg catcgccggg ggtcataacg tgactccctt aattctccgc 6420

tcatgatrag attgtogttt cccgccttca gtttaaaacta tcagtgtttg acaggatata 6480

ttggcgggta aacctaaagc aaaagagcgt ttattagaat aatcggatat ttaaaaggcg 6540

gtgaaaaggc ttatccgttc gtccatttgt atgtgcacgc caaccacagg gttcccccaga 6600

tctggcgccg gccagcgaga cagcaagat tggccgcgcg ccgaaacgat ccgacagcgc 6660

gcccagcaca ggtgcgcagg caaattgcac caacgcatac agcgcacga gaatgccata 6720

gtggcgcggt acgtcgttcg agtgaaccag atcgcgcagg aggcccgcca gcacggcat 6780

aatcaggcgc atgcgcagac cgtcgagcgc gacagtgttc agaattacga tcaggggtat 6840

gttgggttct aogtctggcc tccggacrag cctccgctgg tccattgaa cgcggcgatt 6900

ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt gtcaagcatg 6960

acaaagttag agccgaatac agtgcacgtt gcgcgcctgg acctgttgaa cagagtcggc 7020

gtagacggct tgacgacacg caaactggcg gaacgggttg gggttcacga gccggcgctt 7080

tactggcaat tcaggaaaca gcgggcgctg ctgcagcac tggccgaagc catgctggcg 7140

gagaatcata cgcattcggt gccgagagcc gacgacgact ggcgctcatt tctgatcggg 7200

aatgcccgca gcttcaggca ggcgctgttc gctaccgcg atggcgcgcg catccatgcc 7260

ggcacgcyac cggcgccacc gcagatggaa aoggcgcgcg cgcagcttgg attcctctgc 7320

gaggcggggt tttcgccggg ggcgcgcgtc aatgcgctga tgacaatcag ctacttcaat 7380

gttggggccg tgcttgagga gcaggccggc gacagcatg ccggcgagcg cggcgccacc 7440

gttgaacagg ctccgctctc gccgctgttg cggggccgga tagacgcctt cgacgaagcc 7500

ggtcgggacg cagcgttcga gcagggactc gcggtgattg togatggatt ggcgaaaagg 7560

aggctogttg tcaggaaagt tgaaggacgg agaaaggggtg acgattgac aggaacogctg 7620

cgggagcgca acccaactcac tacagcagag coactgtagac aacatccctt ccccttttcc 7680

acccggtcag acgcccgtag cagcccgcta cgggcttttt catgccttc cctaacctcc 7740

aagcctcaag gccgcgctcg gctctctggt cggcctttct gcgctcttcc gcttctctgc 7800

tcaactgactc gctgcgctcg gtcgttcggc tggggcgagc ggatcagct cactcaaagg 7860

cggtataacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag 7920

gccagcaaaa gcccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 7980

gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcacag 8040

gactataaag ataccaggcg ttccccctg gaagctccct cgtgcgctct cctgttcga 8100

ccctgcgctt taaccgatac ctgtccgoot ttctcccttc ggggaagcgtg gcgcttttcc 8160

gctgcataac cctgcttcgg ggtcattata gcgatttttt cggatatcc atcctttttc 8220

gcacgatata caggattttg ccaaaggggt cgtgtagact ttcttgggtg tatccaaagg 8280

cgtcagccgg gcaggatagg tgaagtaggc ccaccgcga gcgggtgttc cttcttcaact 8340

gtcccttatt cgcacctggc ggtgctcaac gggaaatctg ctctgcgagg ctggccggct 8400

accgcggcgg taacagatga gggcaagcgg atggctgatg aaaccaagcc aaccaggaag 8460

ggcagcccaac ctatcaaggt gtactgctt ccagacgaac gaagagcgat tgaggaaaaag 8520

gcggcgggcg ccggcatgag actgtcggcc taactgctgg ccgtcgccca gggtacaaa 8580

atcacggggc tcgtggacta tgagcagctc cgcgagctgg ccgcaccaa tggcgacctg 8640

ggccgccttg gcgcctgct gaaactctgg ctacccgagc accgcgcac ggcgcggttc 8700

ggatgatgcc cgaatcctcgc cctgctggcg aagatcgaag agaagcagga cgagcttggc 8760

aaggtcatga tgggcgtggt ccgcccgagg gcagagccat gactttttta gccgcataaa 8820

cgcccggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca agaagagcga 8880

cttcgcggag ctggtgaagt acatcaccga cgagcaaggc aagaccgagc gccttttgcga 8940

cgcacacggg gctggttgcg ctgcgcgtg ggttggcgcc cgtctatggc cctgcaaacg 9000

cgccagaaac gccgtogaag cgtgtgcga gacaccggg ccgccggcgt tgtggatacc 9060

tcgcggaana cttggccctc actgacagat gaggggcgga cgttgacact tgaggggccc 9120

actcacccgg cgcgcgcttg acagatgagg ggcaggctcg atttcggcgg gcgacgtgga 9180

gctggccagc ctcgcaaatc ggcgaaaac cctgatttta cgcgagtttc ccacagatga 9240

tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc gcgactactg 9300

acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgcgacaga tgaggggcgc 9360

acctatbgac atttggggg ctgtccaacg gcgaaaac cagcatttgc aagggttttc 9420

gcccgttttt cggccaccgc taaacctgtc tttaacctgc ttttaacca atatttataa 9480

acctgttttt taaccagggc tgcgccctgt gcgcgtgacc gcgcacgcc aaggggggtg 9540

cccccccttc tcgaaccctc ccggcccgtc aaacggggcc tcccatcccc ccaggggctg 9600

cgcccttcgg ccgcgaacgg cctcacccca aaaatggcag cgctggcagt ccttgccatt 9660

gccgggatcg gggcagtaac gggatggcg atcagcccga gcgcgagcc cggaagcatt 9720

gacgtgcgcg aggtgctggc atcgacattc agcgaccagg tgccgggcag tgagggcggc 9780

ggcctgggtg ggggcctgcc cttaacctcg ggcgtcggg cattcacgga cttcatggcg 9840

gggcgcgcaa tttttacctt gggcattctt ggcatagttg tcgcgggtgc cgtgctcgtg 9900

ttcgggggtg cgataaacc agcgaaccat ttgaggtgat aggtaagatt ataccgaggt 9960

atgaaaacga gaattggacc ttacagaat tactctatga agcgccatat ttaaaaagct 10020
accaagacga agaggatgaa gaggatgagg aggcagattg ccttgaatat attgacaata 10080
ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga tttcaggggg 10140
caaggcatag gcagcgcgct tatcaatata tctatagaat gggcaaagca taaaaacttg 10200
catggactaa tgcctgaac ccaggacaat aaccttatag ctgtaaatt ctatcataat 10260
tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgcc gatgaacttg 10320
tcatgcagct ccacogattt tgagaacgac agcgacttcc gtccagcgcg tgccagggtgc 10380
tgctcagat tcaggttatg cgcotcaatt cgtgcgctat atcgttctgt gattacgtgc 10440
agctttccct tcaggcggga ttcatcacgc ggccagccat cgtcatcca tatccaacg 10500
tcaaaggggtg acagcaggct cataagacgc ccagcgctcg ccatagtgcg ttcaccgaat 10560
acgtgcgcga caacgctctt ccggagaactg tcatacgggt aaaacagcca gcgctggcgc 10620
gatttagccc cgcacatagc ccaactgttcg tccatttccg cgcacacgat gacgtcactg 10680
cccggtgta tgcgcgaggt taaccactgc ggctgagtt ttttaagtga cgtaaaatcg 10740
tgttgaggcc aacgccata atgcgggctg ttgccggga tccaacgcca ttcattggcca 10800
tatcaatgat tttctgggtc gtaccgggtt gagaagcgggt gtaagtgaac tgcagttgcc 10860
atgttttacg gcagtgcagc cagagatagc gctgatgtcc ggcggtgctt ttgccgttac 10920
gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa gccactggag 10980
cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc cataattgtg 11040
gtttcaaaat cggctccgto gatactatgt tatacgccaa ctttgaaaaa aactttgaaa 11100
aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg gagttcgtct 11160
tgttataatt agctctctgg ggtatcttta aatactgtag aaaagaggaa ggaataata 11220

aatggctaaa atgagaatat cacoggaatt gaaaaaactg atcgaaaaat accgctgcgt 11280
aaaagatacgg gaaggaatgt ctctgcttaa ggtatataag ctgggtgggag aaaatgaaaa 11340
cctatattta aaaatgacgg acagccggta taaagggacc acctatgatg tggaaacggga 11400
aaaggacatg atgctatggc tggaaaggaaa gctgectggt ccaaaggctc tgcactttga 11460
acggcatgat ggtggagaa atctgctcat gagtgggac gatgcctcc ttctctcga 11520
agagtatgaa gatgaacaaa gccctgaaa gattatcgag ctgtatcggg agtgcacag 11580
gctctttcac tccatcgaca tatcggaatt tccctatacg aatagcttag acagccgctt 11640
agccgaattg gattacttac tgaataacga tctggccgat gtggattgag aaaactggga 11700
agaagaoact ccatttaaag atccgcgcga gctgtatgat tttttaaaga cggaaaagcc 11760
cgaagaggaa cttgtctttt cccacggcga cctggggagac agcaacatct ttgtgaaaga 11820
tggcaaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca agtggatga 11880
cattgccttc tgcgtccggt cgtacaggga ggatatggg gaagaaacgt atgtcgagct 11940
attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt atattttact 12000
ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag caggagcgca 12060
cogactcttt ccgcatcaag tgttttggct ctcaggccga ggccacaggc aagtatttgg 12120
gcaaggggtc gctggtattc gtgcagggca agattcggaa taccaggtac gagaaggagc 12180
gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg gacaccaagg 12240
caccaggcgg gtcaaatcag gaataagggc acattgcccc ggctgtagtc ggggcaatcc 12300
cgcaggagay ytgaaatgaat cggacgtttg accggaaggc atacaggcaa gaactgatcg 12360
acgcgggggtt ttccgcgag gatgcgaaa ccacgcgaag ccgcacgcgc atgcgtgcgc 12420
cccgcgaaac cttccagtc gtcggctcga tggtcagca agctacggcc aagatcgagc 12480

gagacagcgt gcaactgggt cccctgccc tgcccgccc atcgccccc gtggagcgtt 12540

cgcgtcgtct cgaacaggag cgggcaggtt tggcgaagtc gatgaccatc gacacgcgag 12600

gaactatgac gaccaagaag cgaaaaaccg cggcgagga cctggcaaaa caggtcagcg 12660

aggccaagca ggcgcgcttg ctgaacaca cgaagcagca gatcaaggaa atgcagcttt 12720

ccttgcttga tattgcgcgc tggccggaca cgaatgcgagc gatgccaaac gacacggccc 12780

gctctgcctt gttcaccacg cgcaacaaga aaatcccgcg cggcgcgctg caaaacaagg 12840

tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgctcag ctgcggggccg 12900

acgatgacga actggtgtgg cagcagggtt tggagtacgc gaagcgacc cctatcgcg 12960

agcogatcac cttaacgttc tacgagcttt gccaggacct gggctggtcg atcaatggcc 13020

ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg atgggcttca 13080

cgtccgacgc cgttgggac ctggaatcgg tgtgcctgct gcaccccttc cgcgtcctgg 13140

acgtggnaa gaaaacgtcc ctttccagg tctgatcga cagggaatc gtctgtctgt 13200

ttgtggcga ccactacacg aaattcatat gggagaagta ccgcaagctg tcgccgacgg 13260

cccgcagcat gttcagactt ttcagctcgc accgggagcc gtaccgcctc aagctggaaa 13320

ccttcgcct catgtcgga tcggattcca ccgcgtgaa gaagtggcgc gacgaggtcg 13380

gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg gtcaatgatg 13440

acctggtgca ttgcaaacgc tagggccttg tggggctcag tcggcgctgg ggttcagcag 13500

ccagcgcttt actggcattt caggacaacg cgggcactgc tcgacgcact tgcctcgttc 13560

agtatcgctc gggacgcacg gcgcgctcta cgaactgcgc ataaacagag gattaaaaatt 13620

gacaattgtg attaaaggct agattcgacg gcttggagcg gcgcagctgc aggatttcgc 13680

cgaatccga ttgtcgcccc tgaagaaagc tcagagatg ttggggtccg ttacgagca 13740

cgaggagaaaa aagcccatgg aggcgttcgc tgaacgggtg cgagatgcgc tggcattccg 13800

cgctacatc gaaggcgaga tcattgggct gtccgtcttc aaacaggagg acggccccaa 13860

ggagcctcac aaggcgcatc tgcgcggcgt ttctgtggag ccgaacacgc gaggccgagg 13920

ggtcgcgggt atgctgctgc gggcggtgcc ggccgggtta ttgctcgtga tgatcgtccg 13980

acagattcca acgggaahct ggtggatgag catcttcac ctcacccac ttaatatctc 14040

gctattctgg agcttgttgt ttatttcggt ctaccgcctg ccggcggggg tcggcgcgac 14100

ggtagggcgt gtgcagcgc tgatggctgt gttcatctct gccgctctgc taggtagccc 14160

gatacgattg atggcggtcc tgggggctat ttgcggaact gcggcgctgg cgtctgttgt 14220

gttgacacca aacgcagcgc tagatcctgt ccggctcgca gcgggcctgg ccggggcggt 14280

ttccatggcg ttccgaaccg tgctgaccgc caagtggcaa cctccctgc ctctgctcac 14340

ctttaccgcc tggcaactgg ccggcgaggg actctctctc gttccagtag ctttagtgtt 14400

tgatccgcc atcccgatgc ctacaggaac caatgtctc ggcctggcgt ggtcggcct 14460

gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac tcgaacctac 14520

agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc ccgggagca 14580

tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcggtgag caatggatag 14640

gggagttgat atcgtaacg ttoacttcta aagaatagc gccactcagc ttctcagcg 14700

gctttatcca gcgatttctt attatgtcgg catagtcttc aagatcgaca gccctgcacg 14760

gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg agatgatatt 14820

tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct ccgcgagatc 14880

atccgtgttt caaacccggc agcttagttg ccgttcttcc gaatagcatc ggtaacatga 14940

gcgaagtctg ccgccttaca accgctctcc cgtgacgcc gtccggact gatgggctgc 15000

ctgtatcgag tgggtgatttt gtgccgagct gccggtcggg gagctgttgg ctggctgggt 15060
gcaggatata ttgtgggtga aacaaattga cgcttagaca acttaataac acattggcga 15120
cgtttttaat gtactggggg .ggtttttctt ttcaccagtg agacggggcaa cagctgattg 15180
cccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggg ttgccccagc 15240
aggcgaaaat cctgtttgat ggtgggttccg aaatcgccaa aatcccttat aaatcaaaa 15300
aatagcccg gataggggtt agtgtttgtc cagtttgga caagagtcca ctattaaaga 15360
acgtggactc caacgtcaaa gggcgaaaaa cgtctatca gggcgatggc ccactacgtg 15420
aaccatcacc caaatcaagt tttttggggg cgaggtgccg taaagcacta aatcggaacc 15480
ctaaaggag cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaagg 15540
aagggagaa agcgaaagga gcggggccca ttcaggctgc gcaactgttg ggaaggcgga 15600
tcgggtcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggoga 15660
ttaagttggg taacgccag gttttcccg tcaogaogtt gtaaaacgac ggccagtga 15720
ttcagctcg gtaccoggg 15739

<210> 4

<211> 11611

<212> DNA

<213> Artificial

<220>

<223> Vector

<220>

<221> misc_feature

<222> (227)..(227)

<223> n is a, c, g, or t

<220>

<221> misc_feature
 <222> (318)..(318)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (526)..(526)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (8946)..(8946)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (10028)..(10028)
 <223> n is a, c, g, or t

<400> 4
 agcttgcacg cctgcaggtc gagtggagat gtggagtggg cgcttacaca gtacacgagg 60
 acttctagct agaaagaagc attacctcta aacaagtgtc cctgtgcatt ctgggtaaac 120
 gactcatagg agagttgtaa aaaagtttgc gccggcgtat tgggtgttac ggagcattca 180
 ctaggcaacc atgcattggtt actattgtat accatcttag taggaantga ttctgaggtt 240
 tatacctacg atgaatgtgt gtcctgtagg cttgagagtt caaggaagaa acatgcaatt 300
 attcttgcga acccaggngc tgggtgacga attttcatag tcaagctatc agagtaaga 360
 agaggagcat gtcaaagtac aattagagac aaatatatag tcgcgtggag ccaagagcgg 420
 attcctcagt ctctgtagtc tcttgacgac cgttgatctg cttgatctcg tctcccgaaa 480
 atgaaaatag ctctgctaag ctattcttct ctctgcgcga gcttgaaggc cgttactagg 540
 ttgcagtcga tgcattaatg cattgcagat gagctgtatc tgaagaggtt aaaccgaaa 600
 acgcgtttta ttcttgttga catggagcta ttaaatcact agaaggcact ctttgctgct 660

tggaacaatg aacgtatctt atcgagatcc tgaacacccat ttgtctcaac tccggagctg 720

acatcgacac caacgatctt atatccagat tcgtcaagct gtttgatgat ttcagtaacg 780

ttaatggat cgatcccgcg gtcggcatct actctattcc ttggccctcg gacgagtgtc 840

ggggcgctcg ttccactat cggcgagtac ttctacacag ccatcggtcc agacggccgc 900

gcttctgcgg gcgatttgtg tacgcccagc agtcccggt cgggatcgga cgattgcgtc 960

gcacgcagcc tgcgcccaag ctgcacatcc gaaattgccg tcaaccaagc tctgatagag 1020

llygtcaaga ccaatguyga gcatatacgc ccggagccgc ggcatccgt caagctccgg 1080

atgctccgcg tcgaagtagc gcgtctgctg ctccatacaa gccaacccag gcctccagaa 1140

gaagatgttg gcgacctcgt attgggaatc ccgaacatc gcctcgctcc agtcaatgac 1200

cgctgttatg cggccattgt ccgtcaggac attgttggag ccgaatccg cgtgcacgag 1260

gtgccggact tcggggcagt cctcgcccca aagcatcagc tcatcgagag cctgcgcgac 1320

ggaogcactg acggtgtcgt ccatcacagt ttgcagtgta tacacatggg gatcagcaat 1380

cgcgcatatg aaatcacgc atgtagtgtt ttgaccgatt ccttgcggtc cgaatgggac 1440

gaaccgcgtc gtctggctaa gatcgccgcg agcgatcgca tccatggcct ccgcgaccgg 1500

ctgcagaaca gcgggcagtt cggtttcagg caggctcttc aacgtgacac cctgtgcacg 1560

gcgggagatg caatagggtc ggetctcgtc gaattcccca atgtcaagca ctcccggaat 1620

cgggagcgcg gcgatgcaa agtgccgata aacataacga tctttgtaga aaccatcggc 1680

gcagctatct acccgagga catatccagc ccctctaca tcgaagctga aagcacgaga 1740

ttcttcgccc tccgagagct gcacaggtc ggagacgctg tcgaactttt cgatcagaaa 1800

cttctcgaca gacgtcgcg tgagttcagg catggtgatg tctgtctcaag cggggtagct 1860

gttagtcaag ctgcgatgaa gtgggaaagc tcgaactgaa aggttcaaag gaataagggg 1920

tgggaaggat ggagtatgga tgtagcaaag tacttactta ggggaaataa aggttcttgg 1980

atgggaagat gaataactg aagatgggaa aagaaagaga aaagaaaaga gcagctgggtg 2040

gggagagcag gaaaatatg caacaaatgt tggactgacg caacgacctt gtcaaccccg 2100

ccgacacacc gggcgagacg acggggcaaa gctgcctacc agggactgag ggacctcagc 2160

aggtcgagtg cagagcaccg gatgggtcga ctgccagctt gtgttcccgg tctgogccgc 2220

tggccagctc ctgagcggcc ttcccggttt catacccg gcaaacgagg agaggcacga 2280

tatttggacg ccctacagat gccggtggy ccaallaggy agcttacgcy ccgggtactc 2340

gctctaccta cttcggagaa ggtactatct cgtgaatott ttaccagatc ggaagcaatt 2400

ggactctctg acctaggta atggcatgct atttcgcga cggctataca ccctggctt 2460

cacattctcc ttccttact gccggtgatt cgatgaagct ccatattctc cgatgatgca 2520

atagattott ggtcaacgag gggcacacca gcctttccac ttcggggcgg agggggcgcc 2580

ggtcccgat taataatcat ccaactgcacc tcagagccgc cagagctgtc tggccagtgg 2640

cttattactc agccctctc tctgcgtccg tccgtctctc cgcattgccag aaagagtcac 2700

cggtcactgt acagagctca cgagttctc acatttttct acaaatgggtg gagggcgcg 2760

attttaggct caagtcatga cctctgggt cactccagaa tcagctagggt caacgaataa 2820

ggatgattct ataggaagat ccaggcaccg gtcaaccatg atctggacag atttgggagc 2880

tccgtataag ctctccacct atcttattct gtatagttta ggcttaaagt ttatccagga 2940

gatgttgctg aagtcgattt ggtccactt cctcactgggt agctatacga ctttgatgg 3000

cggtgtaggg gctgtattg gtctcgatca aacacaaata gaattaaatg gtactcgagt 3060

ccactgaagg tggcttctcc gtcttcgta gccgtgccga aatccttaca gcttggttg 3120

tgtgactttt ggttacgccg tctgactttt gtggtgagct aactagagat catgctatat 3180

ctcctgattt aatacaatgc tcatcataac attccacctg gaactgctag caacgtttga 3240

cttgcaattgt gcaacgcctt ttgcagagct atcggatgat caatagtgcc acgttctaaa 3300

ttcaaccaac gcagggtgcc caagccttcg acatccggat gtatttcgaa aacctcatgg 3360

cgattgcagt ctcagattc atgttcattc caatgctcat tggagaataa aaggttcaca 3420

gggaataagt tcaaaactcg gatacttgag aatattgaaa gccaaagac cctctatgct 3480

ccaagctaga gtctcagcct ggaaagcaaa tccaaatgaa gctatgctac ctccaattcc 3540

tcatcatcct atctataata cagagtcgaa gaatactctc ttgacaacgc tcogtctccc 3600

gacttcaata aggagcttac tctctcttga caccaccctc ccagttcttc tcggcgctct 3660

ggaggggagg cttgtcggtc ttgggctggc cctggctgag aaagctgttg gcagccttaa 3720

agggagcctg gaggtcacca gtcgctggct tcccgaagac gtggatctta accagattcg 3780

aaagcgccct cagcggatga tcgactggat cagaagagcg ttggtgtact tgaagtacag 3840

atgcatgagc gccatcatgc caacgcccac gaactggctc ttaatgagct ggcggaactg 3900

gcccttatcg tactccatgt tggtagttgt gacaggacga ggctcctcgc cgttccaag 3960

cggagcaggc tcgacgtatt tcagtgtcga aagatctgat caagagacag gatgaggatc 4020

gtttcgcagt attgaacaag atggattgca cgcaggttct ccggccgctt ggggtgagag 4080

gctattcggc tatgactggg cacaacagac aatcggctgc tctgatgcgc ccgtgttccg 4140

gctgtcagc cagggggccc cggttctttt tgtcaagacc gaactgtcgc gtgccctgaa 4200

tgaactgcag gacgaggcag cgcggctatc gtggctggcc acgacgggag ttcttgcgc 4260

agctgtgctc gacgttgtca ctgaagcggg aagggaactg ctgctattgg gcgaagtgcc 4320

ggggcaggat ctcctgtcat ctcaccttgc tctgcgcgag aaagtatcca tcatggctga 4380

tgcaatgcgg cggctgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa 4440

acatcgcatc gagcgagcac gtactcggat ggaagcoggt ctgtcgatc aggatgatct 4500

ggacgaagag catcaggggc tcgcgccagc cgaactgttc gccaggctca aggcgcgcac 4560

gcccgacggc gaggatctcg tcgtgaccca tggcgatgac tgcctgcoga atatcatggt 4620

ggaaaatggc cgcttttctg gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta 4680

tcaggacata gcgttggcta cccgtgatat tgcctgaagag cttggcggcg aatgggctga 4740

ccgcttcctc gtgctttacg gtatcgccgc tcccgattcg cagcgcatcg ccttcatatg 4800

ccttcctgac gagttcttct gagcgggact ctggggttcg aatgaccga ccaagcgacg 4860

cccaacctgc catcacaga ttctogattcc accgccgctc tctatgaaag gttgggcttc 4920

ggaatcgttt tccgggacgc cggctggatg atcctccagc gcggggatct catgctggag 4980

ttcttcgccc acccggggct cgatcccttc gcgagttggt tcagctgctg cctgaggctg 5040

gacgacctcg cggagttcta ccggcagtgc aaatccgtcg gcatacagga aacacagcgc 5100

ggctatccgc gcatacatgc ccccgaaactg caggagtggg gaggcacgat ggcgccttgc 5160

gtccggatct ttgtgaagga accttacttc tgtggtgtga cataattgga caaactacct 5220

acagagattt aaagctctaa ggtaaatata aaatttttaa gtgtataatg tgttaacta 5280

ctgattctaa ttgttttgtt attttagatt ccaacctatg gaactgatga atgggagcag 5340

tgggtgaatg cctttaatga ggaaaacctg ttttgctcag aagaaatgcc atotagtgat 5400

gatgaggcta ctgtgactc tcaacattct actcctccaa aaaagaagag aaaggtagaa 5460

gaccccaagg actttcttc agaatgtcta agllllltga gtcagtgtgt gtttagtaat 5520

agaactcttg ctgtcttgc tatttacacc acaaaggaaa aagctgacct gctatacaag 5580

aaaattatgg aaaaatatc tgtaaccttt ataagtaggc ataacagtta taatcataac 5640

atactgtttt ttcttactcc acacaggcat agagtgtctg ctattaataa ctatgctcaa 5700
 aaattgtgta cctttagcct ttttaattgt aaaggggtta ataaggaata tttgatgtat 5760
 agtgccttga cttagagbca taatcagcca taaccatctt gtagaggttt tacttgcctt 5820
 aaaaaacctc ccacacctcc ccctgaacct gaaacataaa atgaatgcaa ttgttgttgt 5880
 taactgtttt attgragctt aaaaagggtta caataaagc aatagcatca caaatttcac 5940
 aaataaagca tttttttcac tgcattctag ttgtgggttg tcaaaactca tcaatgtatc 6000
 ttatcatgtc tggatctgac gggtcgcgat gatcgtgtc ctgtcgttga ggaccggct 6060
 aggtcggcgg ggttgcctta ctggttagca gaatgaatca ccgatacgcg agcgaacgtg 6120
 aagcgactgc tgcctgaaaa cgtctgcgac ctgagcaaca acatgaatgg tcttcggttt 6180
 ccgtgtttcg taaagtctgg aaacgcggaa gtcagcgtc ttccgcttcc togtcactg 6240
 actcgtcgcg ctgcgtctgt cggctgcggc gagcgggtatc agctcactca aaagcggtta 6300
 tacggttata ccacagatca ggggataaag caggaaagaa catgtgagca aaagccagc 6360
 aabaggccag caaaaggcca ggaaccgtta aaaggccgcg ttgctggcgt ttttccatag 6420
 gctccgcccc cctgaacgag atcacaaaaa tcgacgtca agtcagaggt ggcgaaaccc 6480
 gacaggacta taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcgtgt 6540
 tcgacccctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct 6600
 ttctcatagc tcacgtctga ggtatctcag ttccgtgtag gtcgttcgtc ccaagctggg 6660
 ctgtgtgcac gaaccccccg ttacgcccga ccgtgcgcc ttatccggta actatcgtct 6720
 tgagtccaac ccggttaaac acgacttata gccactggca gcagccactg gtaacaggat 6780
 tagcagagcg aggtatgtag gcggtgtac agagtctctg aagtgggtgc ctaactacgg 6840
 ctacactaga aggacagtat ttggtatctg cgtctgtctg aagccagtta ccttcggaaa 6900

aagagttggt agctcttgat ccggcaaaaca aaccaccgct ggtagcggtg gtttttttgt 6960
ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 7020
tacgggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggatttttg tcattgagatt 7080
atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 7140
aagtatatat gagttaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcaacctat 7200
ctcagcgatc tgtctatttc gttcatccat agttgcctga ctccccctcg tgtagataac 7260
tacgatacgg gagggcttac catctggccc cagtgtctga atgataccgc gagaccacag 7320
ctcacgggct ccagatttat cagcaataaa ccagccagcc ggaagggccg aggcagagaag 7380
tggctctgca actttatccg cctccatcca gtctattaat tggtgccggg aagctagagt 7440
aagtagttcg ccagttaata gtttcgcaa cgttgttgcc attgtctcag gcactgtggt 7500
gtcagcctcg tcgttttgta tggcttcatt cagctccggt tcccaacgat caagcgaggt 7560
taccatgacc cccatgttgt gcaaaaaagc ggttagcttc ttccgtcttc ccatcgttgt 7620
cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct 7680
tactgtcatg ccactcgtaa gatcttttc tgtacttgtt gactactcaa ccaagtcatt 7740
ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg gcgtcaacac gggataatac 7800
cgccccacat agcagaactt taaaagtgtc catcattgga aaacgttctt cggggcgaaa 7860
actctcaagg atcttaccgc tgttgagatc cagttogatg taaccactc gtgcacccaa 7920
ctgatcttca gcattcttta ctctcaccag cgtttctggg tgagcaaaaa caggaaggca 7980
aaatgcgcga aaaaagggaa taaggcgac ccgggaatgt tgaatactca tactcttctt 8040
ttttcaatat tattgaagca ttatcaggg ttattgtctc atgagcggat acatatttga 8100
atgtatttag aaaaataaac aaataggggt tccgcgcaca ttccccgaa aagtgcacc 8160

tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggg gtatcacgag 8220

gccctttcgt cttcaagaat tcgcgccgcg aattaaccct cactaaagga tccctatagt 8280

gagtcgtatt atcgggccgc gaattctcat gtttgaccgc ttatcatcga taagctctgc 8340

ttttgttga cttccattgt tcattccacg gacaaaaaca gagaaaggaa acgacagagg 8400

ccaaaaagct cgtttcaac acctgtcgtt tcctttcttt tcagagggta ttttaataa 8460

aaacattaag ttatgacgaa gaagaacgga aacgccttaa accgaaaaat tttcataaat 8520

agogaaaacc cgcgaggtcg ccgccccgta acaaggcgga tcgccggaag ggaccgcgaa 8580

atgataataa ttatcaattg catactatcg accggcactgc tgcagagata caccacoggg 8640

gaaacattcc atcatgatgg ccgtgcggac ataggaagcc agttcatcca tcgctttctt 8700

gtctgtctgc atttgcttgg tgacatccag cgcgcacat tcagcagcgt ttttcagcgc 8760

gttttcgac aacgtttcaa tgttggtatc aacaccaggt ttaactttga acttatcggc 8820

actgacggtt acctgttctt gcgctggctc atcacgcagg ataccaaggc tgatgttgta 8880

gatattggtc accggctgag ggttttcgat tgcgcctgcg tggatagcac catttgcgat 8940

caggcngtcc ttgatgaatg acactccatt gcgaataagt togaaggaga cgggtgcacg 9000

aatgcgctgg tccagctcgg tcgattgcoct tttgtgcagc agaggatatc atctcaacgc 9060

caaggctcat cgaagcgcaa tattgtctgt caccaaaacg cgtattgacc aggtgttcaa 9120

cggcgaattt ctgccctctt gatgtcagaa aggcaaagtg attttctttc tggttattcag 9180

ttgctgtgtg tcggtttcag caaaaccaag ctgcgcgaat tcggctgtgc agatttagaa 9240

gycagatcac cagacagcaa cggccaacgg aaaacagcgc atacagaaca tccgtcgccg 9300

cgccgacaac gtgataattt ttatgacca tgatttattt ctttttagac gtgagcctgt 9360

cgcacagcaa agccgcgcaa agttcctcga agctagcttc agacgtgtct agatagctct 9420

gctttttgtt gacttccatt gttcattcca cggacaaaaa cagagaaagg aaacgacaga 9480

ggccaaaaag ctgcgtttca gcacctgtcg ttctctttct ttccagaggg tatttttaaat 9540

aaaaacatta agttatgacg aagaagaacg gaaacgcctt aaaccggaat attttataaa 9600

atagcgaaaa cccgcgaggt cgcgcgcccc taacaaggcg gatcgccgga aaggacccgc 9660

aatgataaat aattatcaat tgcatactat cgacggcact gctgccagat aacaccaccg 9720

gggaaacatt coactcatgat ggccgtgcgg acataggaag ccagttcacc catcgctttc 9780

ttgtctgtcg ccatttgctt tgtgacatcc agcgcgcgac attcagcagc gtttttcagc 9840

gcgttttoga tcaacgtttc aatgttggtt tcaacaccag gtttaacttt gaacttatcg 9900

gcactgacgg ttaccttggt ctgcgctggc tcatacgcga ggataccaag gctgatgttg 9960

tagatattgg tcacggcgtg aggggttttcg attgccgctg cgtggatagc accattttcg 10020

atcagggmgt ccttgatgaa tgacactcca ttggaataa gttogaagga gacggtgtca 10080

cgaatgcgct ggtccagctc ggtcgattgc cttttgtgca gcagagggtat caatctcaac 10140

gccaaggctc atogaagcgc aatattgctg ctcacaaaaa cgcgtattga ccaggtgttc 10200

aacggcaaat ttctgccctt ctgatgtcag aaaggcaaag tgattttctt tctgggtattc 10260

agttgctgtg tgcggttttc agcaaaaacca agctcgcgca attcggctgt gcagatttag 10320

aaggcagatc accagacagc aacggccaac ggaaaacagc gcatacagaa catccgtcgc 10380

cgcgcgcgaca acgtgataat ttttatgacc catgatttat ttctttttag acgtgagcct 10440

gtcgcacagc aaagccgcgc aaagtctctc gaccgatgcc cttgagagcc ttcaaccag 10500

tcagctcctt ccggtgggcg cggggcatga ctatcgtcgc cgcacttatg actgtcttct 10560

ttatcatgca actcgtagga caggtgccgg cagcgcctcg ggtcattttc ggcgaggacc 10620

gctttcgcgt gagegcgacg atgatcggcc tgcgccttgc ggtatttcgga atcttgcaag 10680

ccctcgctca agccttcgtc actggtcccg ccaccaaacg ttctggcgag aagcaggcca 10740
ttatcgccgg catggcgccg gacgcgctgg gctacgtctt gctggcgctc gcgacgcgag 10800
gctggatggc ctcccccatt atgattcttc tcgcttcggg cggcatcggg atgcocgcgt 10860
tgcaggccat gctgtccagg caggtagatg acgaccatca gggacagctt caaggatcgc 10920
tcggcgctct taccagccta acttcgatca ttggaccgct gatcgtcacg gcgattttatg 10980
ccgcctcggc gagcacatgg aacgggttgg catggattgt aggcgcgcgc ctataccttg 11040
tctgcctccc cgcgttgct cgcggtgcat ggagccgggc caccctgacc tgaatggaag 11100
ccggcgccac ctgcctaagc gattcaccac tccaagaatt ggagccaato aattcttgcg 11160
gagaactgtg aatgcgcaaa ccaacccttg gcagaacata tccatcgct cgcgatctc 11220
cagcagccgc acgcgggcga tctcgggcag cgttgggtcc tgcagatcgg gctgtggaat 11280
gtgtgtcagt taggtgtgg aaagtccca ggctcccag caggcagaag tatgcaaagc 11340
atgcattcga attagtgcgc aaccagggtg gaaagtccc caggctccc agcaggcaga 11400
agtatgcaaa gcattgatct caattagtca gcaaccatag tccgccccct aactcgcgc 11460
atcccgcccc taactcgcgc cagttccgc cattctcgc ccatggctg actaattttt 11520
tttatttatg cagaggcgga ggcgcctcg gcctctgagc tattccagaa gtagtgagga 11580
ggcttttttg gaggcctagg cttttgcaaa a 11611

<210> 5

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 5
cgatgttagga gggcgtggat a 21

<210> 6
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 6
gcctctgcgg gcgatttg t 21

<210> 7
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 7
tgagaatato accggaattg 20

<210> 8
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 8
agctcgacat actgttcttc c 21

<210> 9

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 9

gtgaatggaa atccatcgcc tgtc

24

<210> 10

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 10

agtgggtact ctaaaggcca tacc

24

<210> 11

<211> 1771

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (166)..(1155)

<400> 11

ggcacgagct tgcacgcaag tcagcgcgcg caagtcacaa cctgccggtc cacagcctca 60

aataataaag agctcaagcg tttgtgcgcg togaogtggc cagtctgcac tgccttgaa 120

ccgcgagtc cccgcgcgcac tgactgccat agcacagcta gaaga atg cag cta gca 177
Met Gln Leu Ala

gcg aca gta atg ttg gag cag ctt acc gga agc gct gag gca ctc aag	225
Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala Glu Ala Leu Lys	
5 10 15 20	
gag aag gag aag gag gtt gca ggc agc tct gac gtg ttg cgt aca tgg	273
Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val Leu Arg Thr Trp	
25 30 35	
gcg acc cag tac tgg ctt cgg tca gaa gag tca gac gcg gcc cgc ccg	321
Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp Ala Ala Arg Pro	
40 45 50	
gga ctg aag aat gcc tac aag cca cca cct tcc gac aca aag ggc atc	369
Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp Thr Lys Gly Ile	
55 60 65	
aca atg gcg cta cgt gtc atc ggc tcc tgg gcc gca gtg ttc ctc cac	417
Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala Val Phe Leu His	
70 75 80	
gcc att ttt caa atc aag ctt cgg acc tcc ttg gac cag ctg cac tgg	465
Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp Gln Leu His Trp	
85 90 95 100	
ctg ccc gtg tca gat gcc aca gct cag ctg gtt agc ggc acg agc agc	513
Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser Gly Thr Ser Ser	
105 110 115	
ctg ctc gac atc gtc gta gta ttc ttt gtc ctg gag ttc ctg tac aca	561
Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu Phe Leu Tyr Thr	
120 125 130	
ggc ctt ttt atc acc acg cat gat gct atg cat ggc acc atc gcc atg	609
Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly Thr Ile Ala Met	
135 140 145	
aga aac agg cag ctt aat gac ttc ttg ggc aga gta tgc atc tcc ttg	657
Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val Cys Ile Ser Leu	
150 155 160	

tac gcc tgg ttt gat tac aac atg ctg cac cgc aag cat tgg gag cac	705
Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys His Trp Glu His	
165 170 175 180	
cac aac cac act ggc gag gtg ggc aag gac cct gac ttc cac agg gga	753
His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp Phe His Arg Gly	
185 190 195	
aac cct ggc att ctg ccc tgg ttt gcc agc ttc atg tcc agc tac atg	801
Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met	
200 205 210	
tgg atg tgg cag ttt gcg cgc ctc gca tgg tgg acg gtg gtc atg cag	849
Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr Val Val Met Gln	
215 220 225	
ctg ctg ggt gcg cca atg gcg aac ctg ctg gtg ttc atg gcg gcc gcg	897
Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala	
230 235 240	
ccc atc ctg tcc gcc ttc cgc ttg ttc tac ttt ggc acg tac atg ecc	945
Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Met Pro	
245 250 255 260	
cac aag cct gag cct ggc gcc gog tca ggc tct tca cca gcc gtc atg	993
His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser Pro Ala Val Met	
265 270 275	
aac tgg tgg aag tgg cgc act agc cag gcg tcc gac ctg gtc agc ttt	1041
Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp Leu Val Ser Phe	
280 285 290	
ctg acc tgc tac cac ttc gac ctg cac tgg gag cac cac cgc tgg ccc	1089
Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His His Arg Trp Pro	
295 300 305	
ttc gcc ccc tgg tgg gag ctg ccc aac tgc cgc cgc ctg tct ggc cga	1137
Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg Leu Ser Gly Arg	
310 315 320	
ggt ctg gtt cct gcc tag ctggacacac tgcagtgggc cctgctgcca	1185

Gly Leu Val Pro Ala

325

gctgggcatg cagggttggtg caggactggg tgagggtgaaa agctgcaggc gctgctgcgc 1245
 gacacgctgc atgggctacc ctgtgtagct gccgccacta ggggaggggg tttgtagctg 1305
 tcgagcttgc cccatggatg aagctgtgta gtggtgcagg gactacccc acaggccaac 1365
 acccttcgag gagatgtctt gcgtcgggag gagtgttggg cagtgtagat gctatgattg 1425
 tatcttaatg ctgaagcctt taggggagcg acacttagtg ctgggcaggc aacgccctgc 1485
 aagggtgcagg cacaagctag gctggacgag gactcggtag caggcaggtg aagaggtgcg 1545
 ggaggggtggt gccacaccca ctgggcaaga ccattgctgca atgctggcgg tggtgcagtg 1605
 agagctgcgt gattaactgg gctatggatt gtttgagcag tctcacttat tctttgatat 1665
 agatactggt caggcaggtc aggagagtga gtatgaacaa gttgagaggt ggtgcgcgtc 1725
 ccctgcgcctt atgaagctgt aacaataaag tggttcaaaa aaaaaa 1771

<210> 12

<211> 329

<212> PRT

<213> Haematococcus pluvialis

<400> 12

Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala

1

5

10

15

Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val

20

25

30

Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp

35

40

45

Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp
50 55 60

Thr Lys Gly Ile Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala
65 70 75 80

Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp
85 90 95

Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser
100 105 110

Gly Thr Ser Ser Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu
115 120 125

Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly
130 135 140

Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val
145 150 155 160

Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys
165 170 175

His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp
180 185 190

Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met
195 200 205

Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr
 210 215 220

Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe
 225 230 235 240

Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly
 245 250 255

Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser
 260 265 270

Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp
 275 280 285

Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His
 290 295 300

His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg
 305 310 315 320

Leu Ser Gly Arg Gly Leu Val Pro Ala
 325

<210> 13

<211> 1662

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (168)..(1130)

<400> 13
 cggggcaact caagaaattc aacagctgca agcgcgcccc agcctcacag cgccaagtga 60
 gctatcgacg tggttgtgag cgctcgacgt ggtccaactga cgggcctgtg agcctctgcy 120
 ctccgtcttc tgccaaattc cgcgtcgggg cctgctaag tcgaaga atg cac gtc 176
 Met His Val
 1
 gca tcg gca cta atg gtc gag cag'aaa ggc agt gag gca gct gct tcc 224
 Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala Ala Ala Ser
 5 10 15
 agc cca gac gtc ttg aga gcg tgg gcg aca cag tat cac atg cca tcc 272
 Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His Met Pro Ser
 20 25 30 35
 gag tcg tca gac gca gct cgt cct gcg cta aag cac gcc tac aaa cct 320
 Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala Tyr Lys Pro
 40 45 50
 cca gca tct gac gcc aag ggc atc acg atg gcg ctg acc atc att ggc 368
 Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr Ile Val Gly
 55 60 65
 acc tgg acc gca gtg ttt tta cac gca ata ttt caa atc agg cta ccg 416
 Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile Arg Leu Pro
 70 75 80
 aca tcc atg gac cag ctt cac tgg ttg cct gtg tcc gaa gcc aca gcc 464
 Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu Ala Thr Ala
 85 90 95
 cag ctt ttg ggc gga agc agc agc cta ctg cac atc gct gca gtc ttc 512
 Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala Ala Val Phe
 100 105 110 115
 att gta ctt gag ttc ctg tac act ggt cta ttc atc acc aca cat gac 560
 Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp
 120 125 130
 gca atg cat ggc acc ata gct ttg agg cac agg cag etc aat gat etc 608

Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu Asn Asp Leu	
135	140
145	
ctt ggc aac atc tgc ata tca ctg tac gcc tgg ttt gac tac agc atg	656
Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Ser Met	
150	155
160	
ctg cat cgc aag cac tgg gag cac cac aac cat act ggc gaa gtg ggg	704
Leu His Arg Lys His Trp Glu His His Asn His Thr Gly Glu Val Gly	
165	170
175	
aaa gac cct gac ttc cac aag gga aat ccc gcc ctt gtc ccc tgg ttc	752
Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val Pro Trp Phe	
180	185
190	195
gcc agc ttc atg tcc agc tac atg tcc ctg tgg cag ttt gcc cgg ctg	800
Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe Ala Arg Leu	
200	205
210	
gca tgg tgg gca gtg gtg atg caa atg ctg ggg gcg ccc atg gca aat	848
Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro Met Ala Asn	
215	220
225	
ctc cta gtc ttc atg gct gca gcc cca atc ttg tca gca ttc cgc ctc	896
Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu	
230	235
240	
ttc tac ttc ggc act tac ctg cca cac aag cct gag cca gcc cct gca	944
Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro Gly Pro Ala	
245	250
255	
gca gcc tct cag gtg atg gcc tgg ttc agg gcc aag aca agt gag gca	992
Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr Ser Glu Ala	
260	265
270	275
tct gat gtg atg agt ttc ctg aca tgc tac cac ttt gac ctg cac tgg	1040
Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp	
280	285
290	
gag cac cac agg tgg ccc ttt gcc ccc tgg tgg cag ctg ccc cac tgc	1088
Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu Pro His Cys	
295	300
305	

cgc cgc ctg tcc ggg cgt ggc ctg gtg cct gcc ttg gca tga 1130
 Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala
 310 315 320

cctggtccct ccgctggtga ccagcgtct gcacaagagt gtcattgctac aggggtgctgc 1190
 gccagtggtg agcgcagtgc actctcagcc tgtatggggc taccgctgtg ccactgagca 1250
 ctgggcatgc cactgagcac tgggcgtgct actgagcaat gggcgtgcta ctgagcaatg 1310
 ggcgctgtac tgacaatggg cgtgctactg gggctctggca gtggttagga tggagtttga 1370
 tgcattcagt agcgggtggcc aacgtcatgt ggalyytggc agtctgagg ggttttaggc 1430
 gccggcattt gagagggcta agttataaat cgcattgtgc tcatgcgcac atatctgcac 1490
 acagccaggg aaatcccttc gagagtgatt atgggacact tgtattgggt tcgtgtctatt 1550
 gttttattca gcagcagtac ttagtgaggg tgagagcagg gtgggtgagag tggagtgtgt 1610
 gagtatgaac ctggctgacg aggtgaacag cctgtaatga atgactctgt ct 1662

<210> 14

<211> 320

<212> PRT

<213> Haematococcus pluvialis

<400> 14

Met His Val Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala
 1 5 10 15

Ala Ala Ser Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His
 20 25 30

Met Pro Ser Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala
 35 40 45

Tyr Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr
50 55 60

Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile
65 70 75 80

Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu
85 90 95

Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala
100 105 110

Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr
115 120 125

Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu
130 135 140

Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp
145 150 155 160

Tyr Ser Met Leu His Arg Lys His Trp Glu His His Asn His Thr Gly
165 170 175

Glu Val Gly Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val
180 185 190

Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe
195 200 205

Ala Arg Leu Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro

210

215

220

Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala
 225 230 235 240

Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro
 245 250 255

Gly Pro Ala Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr
 260 265 270

Ser Glu Ala Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp
 275 280 285

Leu His Trp Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu
 290 295 300

Pro His Cys Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala
 305 310 315 320

<210> 15

<211> 729

<212> DNA

<213> *Agrobacterium aurantiacum*

<220>

<221> CDS

<222> (1)..(729)

<400> 15

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc acc agc ctg
 Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
 1 5 10 15

48

40/358

atc gtc tgc ggc ggc atc atc gcc gct tgg ctg gcc ctg cat gtg cat	96
Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His	
20 25 30	
gcg ctg tgg ttt ctg gac gca gcg gcg cat ccc atc ctg gcg atc gca	144
Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala	
35 40 45	
aat ttc ctg ggg ctg acc tgg ctg tgc gtc gga ttg ttc atc atc gcg	192
Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala	
50 55 60	
cat gac gcg atg cac ggg tgc gtg gtg ccg ggg cgt ccg cgc gcc aat	240
His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn	
65 70 75 80	
gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tgc tgg	288
Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp	
85 90 95	
cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc	336
Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr	
100 105 110	
gac gac gac ccc gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc	384
Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala	
115 120 125	
cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc	432
Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro	
130 135 140	
gtc atc gtg acg gtc tat gcg ctg atc ctt ggg gat cgc tgg atg tac	480
Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr	
145 150 155 160	
gtg gtc ttc tgg ccg ctg ccg tgc atc ctg gcg tgc atc cag ctg ttc	528
Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe	
165 170 175	
gtg ttc ggc acc tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg	576
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro	

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
 65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
 85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr
 100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
 115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
 130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
 145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
 165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
 180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
 195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
 210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
 225 230 235 240

Thr Ala

<210> 17

<211> 1631

<212> DNA

<213> *Alcaligenes* sp.

<220>

<221> CDS

<222> (99)..(827)

<400> 17

ctgcaggcgcg ggcccgggtgg ccaatgggtcg caaccggcag gactggaaca ggacggcggg 60

coggtctagg ctgtcgccct acgcagcagg agtttcgg atg tcc gga cgg aag cct 116
 Met Ser Gly Arg Lys Pro
 1 5

ggc aca act ggc gac acg atc gtc aat ctc ggt ctg acc gcc gcg atc 164
 Gly Thr Thr Gly Asp Thr Ile Val Asn Leu Gly Leu Thr Ala Ala Ile
 10 15 20

ctg ctg tgc tgg ctg gtc ctg cac gcc ttt acg cta tgg ttg cta gat 212
 Leu Leu Cys Trp Leu Val Leu His Ala Phe Thr Leu Trp Leu Leu Asp
 25 30 35

acg gcc gcg cat ccg ctg ctt gcc gtg ctg tgc ctg gct ggg ctg acc 260
 Ala Ala Ala His Pro Leu Leu Ala Val Leu Cys Leu Ala Gly Leu Thr
 40 45 50

tgg ctg tgc gtc ggg ctg ttc atc atc gcg cat gac gca atg cac ggg 308
 Trp Leu Ser Val Gly Leu Phe Ile Ile Ala His Asp Ala Met His Gly
 55 60 65 70

tcc gtg gtg ccg ggg cgg ccg cgc gcc aat gcg gcg atc ggg caa ctg 356
 Ser Val Val Pro Gly Arg Pro Arg Ala Asn Ala Ala Ile Gly Gln Leu

75	80	85	
gcg ctg tgg ctc tat gcg ggg ttc tgc tgg ccc aag ctg atc gcc aag			404
Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp Pro Lys Leu Ile Ala Lys			
90	95	100	
cac atg acg cat cac cgg cac gcc ggc acc gac aac gat ccc gat ttc			452
His Met Thr His His Arg His Ala Gly Thr Asp Asn Asp Pro Asp Phe			
105	110	115	
ggt cac gga ggg ccc gtg cgc tgg tac ggc agc ttc gtc tcc acc tat			500
Gly His Gly Gly Pro Val Arg Trp Tyr Gly Ser Phe Val Ser Thr Tyr			
120	125	130	
ttc ggc tgg cga gag gga ctg ctg cta cgc gtg atc gtc acc acc tat			548
Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro Val Ile Val Thr Thr Tyr			
135	140	145	150
gcg ctg atc ctg ggc gat cgc tgg atg tat gtc atc ttc tgg cgc gtc			596
Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr Val Ile Phe Trp Pro Val			
155	160	165	
cgc gcc gtt ctg gcg tgc atc cag att ttc gtc ttc gga act tgg ctg			644
Pro Ala Val Leu Ala Ser Ile Gln Ile Phe Val Phe Gly Thr Trp Leu			
170	175	180	
ccc cac cgc cgc gga cat gac gat ttt ccc gac cgc cac aac gcg agg			692
Pro His Arg Pro Gly His Asp Asp Phe Pro Asp Arg His Asn Ala Arg			
185	190	195	
tcg acc ggc atc ggc gac cgc ttg tca cta ctg acc tgc ttc cat ttc			740
Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu Leu Thr Cys Phe His Phe			
200	205	210	
ggc gcc tat cac cac gaa cat cac ctg cat ccg cat gtg ccg tgg tgg			788
Gly Gly Tyr His His Glu His His Leu His Pro His Val Pro Trp Trp			
215	220	225	230
cgc ctg cct cgt aca cgc aag acc gga ggc cgc gca tga cgcaattcct			837
Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly Arg Ala			
235	240		

cattgtcgtg gcgacagtcg tcgtgatgga gctgaccgcc tattccgtcc accgctggat 897
 tatgcacggc cccctagggt ggggctggca caagtcccat cagcaagagc acgaccacgc 957
 gttggagaag aacgacctct acggcgtcgt cttcgcggtg ctggcgacga tcctcttcac 1017
 cgtggcgccc tatttggtgc cggtgctgtg gtggatcgcc ctgggcataa cggctctatg 1077
 gttgatctat ttcatctcgc acgacggggt tgtgcatcaa cgctggccgt ttoggtatat 1137
 tcgcggcggg ggcattttcc gcaggctcta ccaagctcat cgctcgacc accggtcga 1197
 gggcggggac cactgcgtca gcttcggctt catctatgcc ccaccgtgg acaagctgaa 1257
 gcaggatctg aagcgttggg gtgtcctcgc ccccaggac gagcgtccgt cgtgatctct 1317
 gatcccgcg tgccgcctg aaatccgacg tgcgtctggc agggcgcgcc cttgccaaag 1377
 gactgatcgc gctggcgatc cgcaaggcgc ggcccgacct tcgctgctg ctgctggacc 1437
 gtgcgggggg cgcctgggaa gggcataact ggtcctgcca cgacaacgat ttggcgccgc 1497
 actggctgga ccgcctgaag ccgatcaggc gtggcgactg gcccgatcag gaggtgcggt 1557
 tcccagacca ttgcggaagg ctccgggccc gatatggctc gatcgacggg cgggggctga 1617
 tgcgtgcggt gacc 1631

<210> 18

<211> 242

<212> PRT

<213> *Alcaligenes* sp.

<400> 18

Met Ser Gly Arg Lys Pro Gly Thr Thr Gly Asp Thr Ile Val Asn Leu

1

5

10

15

Gly Leu Thr Ala Ala Ile Leu Leu Cys Trp Leu Val Leu His Ala Phe
20 25 30

Thr Leu Trp Leu Leu Asp Ala Ala Ala His Pro Leu Leu Ala Val Leu
35 40 45

Cys Leu Ala Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Ile Gly Gln Leu Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Pro Lys Leu Ile Ala Lys His Met Thr His His Arg His Ala Gly Thr
100 105 110

Asp Asn Asp Pro Asp Phe Gly His Gly Gly Pro Val Arg Trp Tyr Gly
115 120 125

Ser Phe Val Ser Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Thr Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Ile Phe Trp Pro Val Pro Ala Val Leu Ala Ser Ile Gln Ile Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Asp Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu
 195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
 210 215 220

Pro His Val Pro Trp Trp Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly
 225 230 235 240

Arg Ala

<210> 19

<211> 729

<212> DNA

<213> *Paracoccus marcusii*

<220>

<221> CDS

<222> (1)..(729)

<400> 19

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc aca agc ctg 48
 Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
 1 5 10 15

atc gtc tgg ggc ggc atc atc gcc gca tgg ctg gcc ctg cat gtc cat 96
 Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
 20 25 30

gcg ctg tgg ttt ctg gac gcg gcg gcc cat ccc atc ctg gcg gtc gcg 144
 Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala
 35 40 45

aat ttc ctg ggg ctg acc tgg ctg tgg gtc gga ttg ttc atc atc gcg 192

48/358

Asn	Phe	Leu	Gly	Leu	Thr	Trp	Leu	Ser	Val	Gly	Leu	Phe	Ile	Ile	Ala		
50							55					60					
cat	gac	gcg	atg	cac	ggg	tcg	gtc	gtg	ccg	ggg	cgt	ccg	cgc	gcc	aat	240	
His	Asp	Ala	Met	His	Gly	Ser	Val	Val	Pro	Gly	Arg	Pro	Arg	Ala	Asn		
65				70					75					80			
gcg	gcg	atg	ggc	cag	ctt	gtc	ctg	tgg	ctg	tat	gcc	gga	ttt	tcg	tgg	288	
Ala	Ala	Met	Gly	Gln	Leu	Val	Leu	Trp	Leu	Tyr	Ala	Gly	Phe	Ser	Trp		
			85					90					95				
cgc	aag	atg	atc	gtc	aag	cac	atg	gcc	cat	cac	cgc	cat	gcc	gga	acc	336	
Arg	Lys	Met	Ile	Val	Lys	His	Met	Ala	His	His	Arg	His	Ala	Gly	Thr		
			100					105					110				
gac	gac	gac	cca	gat	ttc	gac	cat	ggc	ggc	ccg	gtc	cgc	tgg	tac	gcc	384	
Asp	Asp	Asp	Pro	Asp	Phe	Asp	His	Gly	Gly	Pro	Val	Arg	Trp	Tyr	Ala		
			115					120					125				
cgc	ttc	atc	ggc	acc	tat	ttc	ggc	tgg	cgc	gag	ggg	ctg	ctg	ctg	ccc	432	
Arg	Phe	Ile	Gly	Thr	Tyr	Phe	Gly	Trp	Arg	Glu	Gly	Leu	Leu	Leu	Pro		
			130					135					140				
gtc	atc	gtg	acg	gtc	tat	gcg	ctg	atc	ctg	ggg	gat	cgc	tgg	atg	tac	480	
Val	Ile	Val	Thr	Val	Tyr	Ala	Leu	Ile	Leu	Gly	Asp	Arg	Trp	Met	Tyr		
			145				150					155			160		
gtg	gtc	ttc	tgg	ccg	ttg	ccg	tcg	atc	ctg	gcg	tcg	atc	cag	ctg	ttc	528	
Val	Val	Phe	Trp	Pro	Leu	Pro	Ser	Ile	Leu	Ala	Ser	Ile	Gln	Leu	Phe		
				165					170					175			
gtg	ttc	ggc	act	tgg	ctg	ccg	cac	cgc	ccc	ggc	cac	gac	gcg	ttc	ccg	576	
Val	Phe	Gly	Thr	Trp	Leu	Pro	His	Arg	Pro	Gly	His	Asp	Ala	Phe	Pro		
			180					185					190				
gac	cgc	cat	aat	gcg	cgg	tcg	tcg	cgg	atc	agc	gac	cct	gtg	tcg	ctg	624	
Asp	Arg	His	Asn	Ala	Arg	Ser	Ser	Arg	Ile	Ser	Asp	Pro	Val	Ser	Leu		
			195					200					205				
ctg	acc	tgc	ttt	cat	ttt	ggc	ggt	tat	cat	cac	gaa	cac	cac	ctg	cac	672	
Leu	Thr	Cys	Phe	His	Phe	Gly	Gly	Tyr	His	His	Glu	His	His	Leu	His		
			210					215					220				

ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac 720
Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
225 230 235 240

acc gca tga 729
Thr Ala

<210> 20
<211> 242
<212> PRT
<213> *Paracoccus marcusii*

<400> 20

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
1 5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala
35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr

100

105

110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
 115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
 130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
 145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
 165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
 180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
 195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
 210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
 225 230 235 240

Thr Ala

<210> 21

<211> 1629

<212> DNA

<213> *Synechocystis* sp.

<220>

<221> CDS

<222> (1)..(1629)

<400> 21

atg atc acc acc gat gtt gtc att att ggg gcg ggg cac aat ggc tta	48
Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu	
1 5 10 15	

gtc tgt gca gcc tat ttg ctc caa cgg ggc ttg ggg gtg acg tta cta	96
Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu	
20 25 30	

gaa aag cgg gaa gta cca ggg ggg gcg gcc acc aca gaa gct ctc atg	144
Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met	
35 40 45	

cgg gag cta tcc ccc cag ttt cgc ttt aac cgc tgt gcc att gac cac	192
Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His	
50 55 60	

gaa ttt atc ttt ctg ggg cgg gtg ttg cag gag cta aat tta gcc cag	240
Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln	
65 70 75 80	

tat ggt ttg gaa tat tta ttt tgt gac ccc agt gtt ttt tgt cgg ggg	288
Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly	
85 90 95	

ctg gat ggc caa gct ttt atg agc tac cgt tcc cta gaa aaa acc tgt	336
Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys	
100 105 110	

gcc cac att gcc acc tat agc ccc cga gat gcg gaa aaa tat cgg caa	384
Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln	
115 120 125	

ttt gtc aat tat tgg acg gat ttg ctc aac gct gtc cag cct gct ttt	432
---	-----

Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe	
130 135 140	
aat gct ccg ccc cag gct tta cta gat tta gcc ctg aac tat ggt tgg	480
Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp	
145 150 155 160	
gaa aac tta aaa tcc gtg ctg gcg atc gcc ggg tcg aaa acc aag gcg	528
Gln Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala	
165 170 175	
ttg gat ttt atc cgc act atg atc ggc tcc ccg gaa gat gtg ctc aat	576
Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn	
180 185 190	
gaa tgg ttc gac agc gaa cgg gtt aaa gct cct tta gct aga cta tgt	624
Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys	
195 200 205	
tcg gaa att ggc gct ccc cca tcc caa aag ggt agt agc tcc ggc atg	672
Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met	
210 215 220	
atg atg gtg gcc atg cgg cat ttg gag gga att gcc aga cca aaa gga	720
Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly	
225 230 235 240	
ggc act gga gcc ctc aca gaa gcc ttg gtg aag tta gtg caa gcc caa	768
Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln	
245 250 255	
ggg gga aaa atc ctc act gac caa acc gtc aaa cgg gta ttg gtg gaa	816
Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu	
260 265 270	
aac aac cag gcg atc ggg gtg gag gta gct aac gga gaa cag tac cgg	864
Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg	
275 280 285	
gcc aaa aaa ggc gtg att tct aac atc gat gcc cgc cgt tta ttt ttg	912
Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu	
290 295 300	

caa ttg gtg gaa cgg ggg gcc cta gcc aag gtg aat caa aac cta ggg Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly 305 310 315 320	960
gaa cga ctg gaa cgg cgc act gtg aac aat aac gaa gcc att tta aaa Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys 325 330 335	1008
atc gat tgt gcc ctc tcc ggt tta ccc cac ttc act gcc atg gcc ggg Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly 340 345 350	1056
cgg gag gat cta acg gga act att ttg att gcc gac tcg gta cgc cat Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His 355 360 365	1104
gtc gag gaa gcc cac gcc ctc att gcc ttg ggg caa att ccc gat gct Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala 370 375 380	1152
aat cgg tct tta tat ttg gat att ccc act gta ttg gac ccc acc atg Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met 385 390 395 400	1200
gcc ccc cct ggg cag cac acc ctc tgg atc gaa ttt ttt gcc ccc tac Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr 405 410 415	1248
cgc atc gcc ggg ttg gaa ggg aca ggg tta atg gcc aca ggt tgg acc Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr 420 425 430	1296
gat gag tta aag gaa aaa gtg gcg gat cgg gtg att gat aaa tta acg Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr 435 440 445	1344
gac tat gcc cct aac cta aaa tct ctg atc att ggt cgc cga gtg gaa Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu 450 455 460	1392
agt ccc gcc gaa ctg gcc caa cgg ctg gga agt tac aac ggc aat gtc	1440

54/358

Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val
 465 470 475 480

tat cat ctg gat atg agt ttg gac caa atg atg ttc ctc cgg cct cta 1488
 Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu
 485 490 495

ccg gaa att gcc aac tac caa acc ccc atc aaa aat ctt tac tta aca 1536
 Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr
 500 505 510

ggg gcg ggt acc cat ccc ggt ggc tcc ata tca ggt atg ccc ggt aga 1584
 Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg
 515 520 525

aat tgc gct cgg gtc ttt tta aaa caa caa cgt cgt ttt tgg taa 1629
 Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp
 530 535 540

<210> 22

<211> 542

<212> PRT

<213> *Synechocystis* sp.

<400> 22

Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu
 1 5 10 15

Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu
 20 25 30

Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met
 35 40 45

Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His
 50 55 60

Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln
65 70 75 80

Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly
85 90 95

Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys
100 105 110

Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln
115 120 125

Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe
130 135 140

Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp
145 150 155 160

Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala
165 170 175

Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn
180 185 190

Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys
195 200 205

Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met
210 215 220

Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly

225 230 235 240

Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln
245 250 255

Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu
260 265 270

Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg
275 280 285

Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu
290 295 300

Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly
305 310 315 320

Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Glu Ala Ile Leu Lys
325 330 335

Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly
340 345 350

Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His
355 360 365

Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala
370 375 380

Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met
385 390 395 400

Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr
405 410 415

Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr
420 425 430

Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr
435 440 445

Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu
450 455 460

Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val
465 470 475 480

Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu
485 490 495

Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr
500 505 510

Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg
515 520 525

Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp
530 535 540

<210> 23

<211> 776

<212> DNA

<213> *Bradyrhizobium* sp.

<220>

<221> CDS

<222> (1)..(774)

<400> 23

atg cat gca gca acc gcc aag gct act gag ttc ggg gcc tct cgg cgc 48

Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg

1 5 10 15

gac gat gcg agg cag cgc cgc gtc ggt ctc acg ctg gcc gcg gtc atc 96

Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile

20 25 30

atc gcc gcc tgg ctg gtg ctg cat gtc ggt ctg atg ttc ttc tgg cgc 144

Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro

35 40 45

ctg acc ctt cac agc ctg ctg ccg gct ttg cct ctg gtg gtg ctg cag 192

Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln

50 55 60

acc tgg ctc tat gta gcc ctg ttc atc atc gcg cat gac tgc atg cac 240

Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His

65 70 75 80

ggc tcg ctg gtg ccg ttc aag ccg cag gtc aac cgc cgt atc gga cag 288

Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln

85 90 95

ctc tgc ctg ttc ctc tat gcc ggg ttc tcc ttc gac gct ctc aat gtc 336

Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val

100 105 110

gag cac cac aag cat cac cgc cat ccc ggc acg gcc gag gat ccc gat 384

Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp

115 120 125

ttc gac gag gtg ccg ccg cac gcc ttc tgg cac tgg ttc gcc agc ttt 432

Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe

130 135 140

ttc ctg cac tat ttc ggc tgg aag cag gtc gcg atc atc gca gcc gtc 480
 Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val
 145 150 155 160

tcg ctg gtt tat cag ctg gtc ttc gcc gtt ccc ttg cag aac atc ctg 528
 Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu
 165 170 175

ctg ttc tgg gcg ctg ccc ggg ctg ctg tcg gcg ctg cag ctg ttc acc 576
 Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr
 180 185 190

ttc ggc acc tat ctg ccg cac aag ccg gcc acg cag ccc ttc gcc gat 624
 Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp
 195 200 205

cgc cac aac gcg cgg acg agc gaa ttt ccc gcg tgg ctg tcg ctg ctg 672
 Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu
 210 215 220

acc tgc ttc cac ttc ggc ttt cat cac gag cat cat ctg cat ccc gat 720
 Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp
 225 230 235 240

gcg ccg tgg tgg cgg ctg ccg gag atc aag cgg cgg gcc ctg gaa agg 768
 Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg
 245 250 255

cgt gac ta 776
 Arg Asp

<210> 24

<211> 258

<212> PRT

<213> Bradyrhizobium sp.

<400> 24

Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg

1	5	10	15
Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile			
	20	25	30
Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro			
	35	40	45
Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln			
	50	55	60
Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His			
	65	70	75
			80
Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln			
	85	90	95
Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val			
	100	105	110
Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp			
	115	120	125
Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe			
	130	135	140
Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val			
	145	150	155
			160
Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu			
	165	170	175

Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr
 180 185 190

Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp
 195 200 205

Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu
 210 215 220

Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp
 225 230 235 240

Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg
 245 250 255

Arg Asp

<210> 25

<211> 777

<212> DNA

<213> Nostoc sp.

<220>

<221> CDS

<222> (1)..(777)

<400> 25

atg gtt cag tgt caa cca tca tct ctg cat tca gaa aaa ctg gtg tta 48
 Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu
 1 5 10 15

ttg tca tgg aca atc aga gat gat aaa aat att aat aag ggt ata ttt 96
 Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe

20	25	30	
att gcc tgc ttt atc tta ttt tta tgg gca att agt tta atc tta tta			144
Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu			
35	40	45	
ctc tca ata gat aca tcc ata att cat aag agc tta tta ggt ata gcc			192
Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala			
50	55	60	
atg ctt tgg cag acc ttc tta tat aca ggt tta ttt att act gct cat			240
Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His			
65	70	75	80
gat gcc atg cac ggc gta gtt tat ccc aaa aat ccc aga ata aat aat			288
Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn			
85	90	95	
ttt ata ggt aag ctc act cta atc ttg tat gga cta ctc cct tat aaa			336
Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys			
100	105	110	
gat tta ttg aaa aaa cat tgg tta cac cac gga cat cct ggt act gat			384
Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp			
115	120	125	
tta gac cct gat tat tac aat ggt cat ccc caa aac ttc ttt ctt tgg			432
Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp			
130	135	140	
tat cta cat ttt atg aag tct tat tgg cga tgg acg caa att ttc gga			480
Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly			
145	150	155	160
tta gtg atg att ttt cat gga ctt aaa aat ctg gtg cat ata cca gaa			528
Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu			
165	170	175	
aat aat tta att ata ttt tgg atg ata cct tct att tta agt tca gta			576
Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val			
180	185	190	

63/358

caa cta ttt tat ttt ggt aca ttt ttg cct cat aaa aag cta gaa ggt 624
 Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly
 195 200 205

ggt tat act aac ccc cat tgt gcg cgc agt atc cca tta cct ctt ttt 672
 Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe
 210 215 220

tgg tct ttt qtt act tgt tat cac ttc ggc tac cac aag gaa cat cac 720
 Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His
 225 230 235 240

gaa tac cct caa ctt cct tgg tgg aaa tta cct gaa gct cac aaa ata 768
 Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile
 245 250 255

tct tta taa 777
 Ser Leu

<210> 26

<211> 258

<212> PRT

<213> Nostoc sp.

<400> 26

Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu
 1 5 10 15

Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe
 20 25 30

Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu
 35 40 45

Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala
 50 55 60

Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His
65 70 75 80

Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn
85 90 95

Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys
100 105 110

Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp
115 120 125

Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp
130 135 140

Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly
145 150 155 160

Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu
165 170 175

Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val
180 185 190

Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly
195 200 205

Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe
210 215 220

Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His
 225 230 235 240

Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile
 245 250 255

Ser Leu

<210> 27

<211> 789

<212> DNA

<213> Nostoc punctiforme

<220>

<221> CDS

<222> (1)..(789)

<400> 27

ttg aat ttt tgt gat aaa cca gtt agc tat tat gtt gca ata gag caa 48
 Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln
 1 5 10 15

tta agt gct aaa gaa gat act gtt tgg ggg ctg gtg att gtc ata gta 96
 Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val
 20 25 30

att att agt ctt tgg gta gct agt ttg gct ttt tta cta gct att aat 144
 Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn
 35 40 45

tat gcc aaa gtc cca att tgg ttg ata cct att gca ata gtt tgg caa 192
 Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln
 50 55 60

atg ttc ctt tat aca ggg cta ttt att act gca cat gat gct atg cat 240
 Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His
 65 70 75 80

ggg tca gtt tat cgt aaa aat ccc aaa att aat aat ttt atc ggt tca	288
Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser	
85 90 95	
cta gct gta gcg ctt tac gct gtg ttt cca tat caa cag atg tta aag	336
Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys	
100 105 110	
aat cat tgc tta cat cat cgt cat cct gct agc gaa gtt gac cca gat	384
Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp	
115 120 125	
ttt cat gat ggt aag aga aca aac gct att ttc tgg tat ctc cat ttc	432
Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe	
130 135 140	
atg ata gaa tac tcc agt tgg caa cag tta ata gta cta act atc cta	480
Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu	
145 150 155 160	
ttt aat tta gct aaa tac gtt ttg cac atc cat caa ata aat ctc atc	528
Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile	
165 170 175	
tta ttt tgg agt att cct cca att tta agt tcc att caa ctg ttt tat	576
Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr	
180 185 190	
ttc gga aca ttt ttg cct cat cga gaa ccc aag aaa gga tat gtt tat	624
Phe Gly Tyr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr	
195 200 205	
ccc cat tgc agc caa aca ata aaa ttg cca act ttt ttg tca ttt atc	672
Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile	
210 215 220	
gct tgc tac cac ttt ggt tat cat gaa gaa cat cat gag tat ccc cat	720
Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His	
225 230 235 240	
gta cct tgg tgg caa ctt cca tct gta tat aag cag aga gta ttc aac	768

Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
 245 250 255

aat tca gta acc aat tgg taa
 Asn Ser Val Thr Asn Ser
 260

789

<210> 28
 <211> 262
 <212> PRT
 <213> Nostoc punctiforme

<400> 28

Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln
 1 5 10 15

Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val
 20 25 30

Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn
 35 40 45

Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln
 50 55 60

Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His
 65 70 75 80

Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser
 85 90 95

Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys
 100 105 110

Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp
 115 120 125

Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe
 130 135 140

Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu
 145 150 155 160

Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile
 165 170 175

Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr
 180 185 190

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr
 195 200 205

Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile
 210 215 220

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
 225 230 235 240

Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
 245 250 255

Asn Ser Val Thr Asn Ser
 260

<211> 762
<212> DNA
<213> Nostoc punctiforme

```
<220>
<221> CDS
<222> (1)..(762)
```

[illegible]

70/358

ttt cac aat ggt aaa cac caa agt ttc ttt gct tgg tat ttt cat ttt 432
 Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe
 130 135 140

atg aaa ggt tac tgg agt tgg ggg caa ata att gcg ttg act att att 480
 Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile
 145 150 155 160

tat aac ttt gct aaa tac ata ctg cat atc cca agt gat aat cta act 528
 Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr
 165 170 175

tac ttt tgg gtg cta ccc tog ctt tta agt tca tta caa tta ttc tat 576
 Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr
 180 185 190

ttt ggt act ttt tta ccc cat agt gaa cca ata ggg ggt tat gtt cag 624
 Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln
 195 200 205

cct cat tgt gcc caa aca att agc cgt cct att tgg tgg tca ttt atc 672
 Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile
 210 215 220

acg tgc tat cat ttt ggc tac cac gag gaa cat cac gaa tat cct cat 720
 Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
 225 230 235 240

att tct tgg tgg cag tta cca gaa att tac aaa gca aaa tag 762
 Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys
 245 250

<210> 30

<211> 253

<212> PRT

<213> Nostoc punctiforme

<400> 30

Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro

1

5

10

15

Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val
20 25 30

Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp
35 40 45

Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln
50 55 60

Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His
65 70 75 80

Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr
85 90 95

Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys
100 105 110

Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp
115 120 125

Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe
130 135 140

Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile
145 150 155 160

Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr
165 170 175

72/358

Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr
 180 185 190

Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln
 195 200 205

Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile
 210 215 220

Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
 225 230 235 240

Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys
 245 250

<210> 31

<211> 1608

<212> DNA

<213> *Haematococcus pluvialis*

<220>

<221> CDS

<222> (3)..(971)

<400> 31

ct aca ttt cac aag ccc gtg agc ggt goa agc gct ctg ccc cac atc 47
 Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile
 1 5 10 15

ggc cca cct cct cat ctc cat cgg tca ttt gct gct acc acg atg ctg 95
 Gly Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu
 20 25 30

tcg aag ctg cag tca atc agc gtc aag gcc cgc cgc gtt gaa cta gcc 143
 Ser Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala
 35 40 45

cgc gac atc acg cgg ccc aaa gtc tgc ctg cat gct cag cgg tgc tcg	191
Arg Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser	
50 55 60	
tta gtt cgg ctg cga gtg gca gca cca cag aca gag gag gcg ctg gga	239
Leu Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly	
65 70 75	
acc gtg cag gct gcc ggc gcg ggc gat gag cac agc gcc gat gta gca	287
Thr Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala	
80 85 90 95	
ctc cag cag ctt gac cgg gct atc gca gag cgt cgt gcc cgg cgc aaa	335
Leu Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys	
100 105 110	
cgg gag cag ctg tca tac cag gct gcc gcc att gca gca tca att ggc	383
Arg Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly	
115 120 125	
gtg tca ggc att gcc atc ttc gcc acc tac ctg aga ttt gcc atg cac	431
Val Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His	
130 135 140	
atg acc gtg ggc ggc gca gtg cca tgg ggt gaa gtg gct ggc act ctc	479
Met Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu	
145 150 155	
ctc ttg gtg gtt ggt ggc gcg ctc ggc atg gag atg tat gcc cgc tat	527
Leu Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr	
160 165 170 175	
gca cac aaa gcc atc tgg cat gag tcg cct ctg ggc tgg ctg ctg cac	575
Ala His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His	
180 185 190	
aag agc cac cac aca cct cgc act gga ccc ttt gaa gcc aac gac ttg	623
Lys Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu	
195 200 205	
ttt gca atc atc aat gga ctg ccc gcc atg ctc ctg tgt aac ttt ggc	671

Phe Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly	
210	215 220
ttc tgg ctg ccc aac gtc ctg ggg gcg gcc tgc ttt gga gcg ggg ctg	719
Phe Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu	
225	230 235
ggc atc acg cta tac ggc atg gca tat atg ttt gta cac gat ggc ctg	767
Gly Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu	
240	245 250 255
gtg cac agg cgc ttt ccc acc ggg ccc atc gct ggc ctg ccc tac atg	815
Val His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met	
	260 265 270
aag cgc ctg aca gtg gcc cac cag cta cac cac agc ggc aag tac ggt	863
Lys Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly	
275	280 285
ggc gcg ccc tgg ggt atg ttc ttg ggt cca cag gag ctg cag cac att	911
Gly Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile	
290	295 300
cca ggt gcg gcg gag gag gtg gag cga ctg gtc ctg gaa ctg gac tgg	959
Pro Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp	
305	310 315
tcc aag cgg tag ggtcgggaac caggcacgct ggttccacac ctcatgectg	1011
Ser Lys Arg	
320	
tgataaggta tgcctagagc gatgcgtgtg agacgggtat gtcacggctg actggctctga	1071
tgcccaatgg catcggccat gtctggctcat cacgggctgg ttgcctgggt gaagggtgatg	1131
cacatcatca tgtgcggttg gaggggctgg cacagtgtgg gctgaactgg agcagttgtc	1191
caggctggcg ttgaatcagt gagggtttgt gattggcggt tgtgaagcaa tgactccgcc	1251
catattctat ttgtgggagc tgagatgatg gcattgcttg gatgtgcatg gatcatggta	1311
gtgcagcaaa ctatattcac ctagggtgt tggtaggatc aggtgaggcc ttgcacattg	1371

catgatgtac tcgtcatggt gtgttggtga gaggatggat gtggatggat gtgtattctc 1431
 agacgtagac cttgactgga ggcttgatcg agagagtggg ccgtattctt tgagaggoga 1491
 ggctcgtgcc agaaatggtg agtggatgac tgtgacgctg tacattgcag gcaggtgaga 1551
 tgcactgtct cgattgtaaa atacattcag atgcaaaaaa aaaaaaaaaa aaaaaaa 1608

<210> 32

<211> 322

<212> PRT

<213> Haematococcus pluvialis

<400> 32

Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile Gly
 1 5 10 15

Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu Ser
 20 25 30

Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala Arg
 35 40 45

Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser Leu
 50 55 60

Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly Thr
 65 70 75 80

Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala Leu
 85 90 95

Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys Arg

100 105 110

Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly Val
115 120 125

Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His Met
130 135 140

Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu Leu
145 150 155 160

Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr Ala
165 170 175

His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His Lys
180 185 190

Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu Phe
195 200 205

Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly Phe
210 215 220

Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu Gly
225 230 235 240

Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu Val
245 250 255

His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met Lys
260 265 270

Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly Gly
 275 280 285

Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile Pro
 290 295 300

Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp Ser
 305 310 315 320

Lys Arg

<210> 33

<211> 528

<212> DNA

<213> Erwinia uredovora

<220>

<221> CDS

<222> (1)..(528)

<400> 33

atg ttg tgg att tgg aat gcc ctg atc gtt ttc gtt acc gtg att ggc 48
 Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly
 1 5 10 15

atg gaa gtg att gct gca ctg gca cac aaa tac atc atg cac ggc tgg 96
 Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp
 20 25 30

ggt tgg gga tgg cat ctt tca cat gaa ccg cgt aaa ggt gcg ttt 144
 Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe
 35 40 45

gaa gtt aac gat ctt tat gcc gtg gtt ttt gct gca tta tcg atc ctg 192
 Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ala Leu Ser Ile Leu

50	55	60	
ctg att tat ctg ggc agt aca gga atg tgg ccg ctc cag tgg att ggc			240
Leu Ile Tyr Leu Gly Ser Thr Gly Met Trp Pro Leu Gln Trp Ile Gly			
65	70	75	80
gca ggt atg acg gcg tat gga tta ctc tat ttt atg gtg cac gac ggg			288
Ala Gly Met Thr Ala Tyr Gly Leu Leu Phe Met Val His Asp Gly			
85	90		95
ctg gtg cat caa cgt tgg cca ttc cgc tat att cca cgc aag ggc tac			336
Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr			
100	105		110
ctc aaa cgg ttg tat atg gcg cac cgt atg cat cac gcc gtc agg ggc			384
Leu Lys Arg Leu Tyr Met Ala His Arg Met His His Ala Val Arg Gly			
115	120		125
aaa gaa ggt tgt gtt tct ttt ggc ttc ctc tat gcg ccg ccc ctg tca			432
Lys Glu Gly Cys Val Ser Phe Gly Phe Leu Tyr Ala Pro Pro Leu Ser			
130	135		140
aaa ott cag gcg aog ctc ogg gaa aga cat ggc gct aga gog ggc gct			480
Lys Leu Gln Ala Thr Leu Arg Glu Arg His Gly Ala Arg Ala Gly Ala			
145	150		155
gcc aga gat gcg cag ggc ggg gag gat gag ccc gca tcc ggg aag taa			528
Ala Arg Asp Ala Gln Gly Gly Glu Asp Glu Pro Ala Ser Gly Lys			
165	170		175
<210> 34			
<211> 175			
<212> PRT			
<213> Erwinia uredovora			
<400> 34			
Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly			
1	5	10	15

79/358

Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp
20 25 30

Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe
35 40 45

Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ala Leu Ser Ile Leu
50 55 60

Leu Ile Tyr Leu Gly Ser Thr Gly Met Trp Pro Leu Gln Trp Ile Gly
65 70 75 80

Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly
85 90 95

Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr
100 105 110

Leu Lys Arg Leu Tyr Met Ala His Arg Met His His Ala Val Arg Gly
115 120 125

Lys Glu Gly Cys Val Ser Phe Gly Phe Leu Tyr Ala Pro Pro Leu Ser
130 135 140

Lys Leu Gln Ala Thr Leu Arg Glu Arg His Gly Ala Arg Ala Gly Ala
145 150 155 160

Ala Arg Asp Ala Gln Gly Gly Glu Asp Glu Pro Ala Ser Gly Lys
165 170 175

<210> 35

<211> 1520

<212> DNA

<213> Artificial

<220>

<223> Promotor

<400> 35

ctcgagtacc gaggcggaac ggcaggaatg ttccctctc ttttagaggg caattcttta 60
tccaatgtca tgttgatgct agatatttct gtctcttata ataaggcgaa taccattttt 120
tgaattgaag ttgagataaa aaaaaagggg gcccaatttg tcaacgcaa agagtcaagc 180
ttttcttttg gctttagcgg aacaatctaa gacttattgt ttttgaagat atttgacctt 240
ttctagatat tccttcaagt aaagcttttt tctagttttt tttttttttc ttgtgtaagg 300
atttatttgt attggtatcc attttttatt ggaagacaag ataagttaat attgattttg 360
cttaaagatt aaaaggaaat cagaaaacga caataaaaa tgtaacggac aaactatggt 420
gtcgattata agtctaaatc cttaaaaaat gacaacgagt tgctttctct tgaaaacaat 480
tcctttgtct ttgcaagaaa ggtttctttt ttgtttgctt gcattactta aacatcaaat 540
caaatgaaag gaataaagca gatttgaggg cgaataagga ttttctggtc aacaagatgt 600
gagtgcaccc taaggaaact aatgccatcc atttgtttta aaacgcacatc aaagattgat 660
gatcaacagg attgagagag agaaaaagaa ctogtgtcat ttatttctgt tgactgaaat 720
tttatattta gaaaaaatgt caaatctata gcttttagcta tattacataa catttgaaat 780
aataataata aaaaaagaca cattagagac acttttcaaa ctctaaataa ctgtctataa 840
acacaaagaa aacaaagacc tctataacaa ctatttagat ttttctogta cttttgtcta 900
aagatgatgt attcttgcta tcccacactt ctttcatttg ttcttgatgc tactaaatat 960
acaaaatttc ttttttgcaa gagatattat tccaaaaatt ttcaaaaaga aatttttttc 1020
acaatagcag ttgatcgtgt aacccaaaga ggttctttgt tattttgcac ttccgctttg 1080

cgggtgatgca tattcaaagt aatatatgga ataaacaacg tgtttaagca tgaagaagaag 1140
gaacaaazagg cgcgtttgaa caaatgcata atatttcaga caaaatgat ctaaagcaag 1200
cagtaaatca aacaagaaac attgctgatt cgcgttagaa aacgataaaa gtctaataag 1260
ccactaagta tacttcaatg aactttttgt atgcttatgg tccaatcaga ccaataattt 1320
gtgaccattc ctgaggtggc tttgggtgatg cggaaacaga aaaaaatttt ctaccacatc 1380
gatttaaaaa acaattttctg ctttgaacca aaactttttt tttctcttta atcattaact 1440
ttatcaagta tgtacctacc ctcaaagtc tcaactcaagc acaattatgc laacallglt 1500
ccaccttctc tttagaaatg 1520

<210> 36

<211> 16245

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 36
 ccgggctggt tgcctctgcc gctgggctgg cggcgtcta tggcctgca aacgcgccag 60
 aaacgcgcgc gaagcgtgt gcgagacacc gcggccgcgc gcgttgtgga tacctcgcgg 120
 aaaacttgcc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180
 ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccgcgcacg tggagctggc 240
 cagcctcgca aatcggcgaa aacgcctgat ttacgcgag ttcccacag atgatgtgga 300
 caagcctggg gataagtgc ctgcgttatt gacacttgag gggcgcgact actgacagat 360
 gagggcgcgc atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat 420
 tgacatttga gggcgtgtcc acaggcagaa aatccagcat ttgcaagggt tcccgcogt 480
 ttttcggcca ccgctaaact gtcttttaac ctgcttttaa accaatattt ataaccttg 540
 tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccc 600
 cttctcgaac cctcccgcc cgctaacgcg ggctcccat cccccaggg gctgcgcgcc 660
 tcggccgcga acggcctcac ccaaaaatg gcagcgttg cagtccttgc cattgccgg 720
 atcggggcag taacgggatg ggcgatoagc ccgagcgcga cgcgcggaag cattgacgtg 780
 ccgcaggtgc tggcatcgac attcagcgac cagggtgcgc gcagtgaggg cggcggcctg 840
 ggtggcggcc tgccttcac ttcggcgcgc ggggcattca cggacttcat ggcggggccg 900
 gcaattttta ccttcggcat tottggcata gtggtgcggg gtgcgtgtct cgtgttcggg 960
 ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020
 acgagaattg gacctttaca gaattactct atgaagcgc atattttaaa agctaccaag 1080
 acgaagagga tgaagaggat gaggggcag attgccttga atatattgac aatactgata 1140
 agataatata tcttttatat agaagatgc gccgtatgta aggatttcag ggggcaaggg 1200
 ataggcagcg cgcttatcaa tatacttata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgcttg aaaccaggga caataacctt atagottgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttacagataat gcccgatgac ttgtcatgc 1380
agctccacog attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgctc 1440
agattcaggt tatgcgctc aatcgctgc gtatatcgt tgctgattac gtgcagcttt 1500
cccttcaggg gggattcata cagcgccag ccacgtca tccatatcac cagctcaaag 1560
ggtgacagca ggtcataag acgcccagc gtgcgatag tgcttccacc gaatacgtgc 1620
gcaacaacgg tcttcggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttctccatt tcggcgaga cgatgaagtc actgccggc 1740
tgtatgcggc aggttacga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttccc ggcatccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgctgat gtccggcggg gcttttgccg ttacgcacca 1980
cccgtcagt agctgaacag gaggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100
aaatcggtc cgtcgatact atgttatag ccaactttga aaacaacttt gaaaaagctg 2160
ttttttgta ttttaagttt taqaatgaa ggaacagtga attggagttc gtctgttat 2220
aattagcttc ttggggtatc ttttaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcacccg aattgaaaaa actgatcgaa aaatacogct gcgtaaaaa 2340
tacggaagga atgtctctg ctaaggata taagctggtg ggagaaaatg aaacctata 2400
tttaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tgggtggaag gaaagctgcc tgttccaaag gtctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcattgagtga ggccgatggc gtcctttgct cggagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtga tcaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggagaaga 2760

cactccattt aaagatccgc gogagctgta tgatttttta aagacggaaa agcccgaga 2820

ggaaactgtc ttttccacg gcgacctggg agacagaaac atctttgtga aagatggcaa 2880

agtaagtggc ttatttgatc ttgggagaag cggcagggcg gacaagtggc atgacattgc 2940

cttctgcgtc cggctgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgcgcgcga caagcaggag cgcaccgact 3120

tcctccgcac caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggctcgtggt attcgtgcag ggcaagattc ggaatacaa gtaacgagaag gacggccaga 3240

cggctacagg gaccgacttc attgcccata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccgcgctg agtcggggca atcccgcaag 3360

gagggatgaat gaatcgagc tttagaccga aggcatacag gcaagaactg atcgacggg 3420

agttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgcccgcg 3480

aaaccttcoa gtccgtcgcc tcgatgggtc ccgaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggtcccccct gccctgcccg cgcacatggc cgcctgggag cgttcgcgtc 3600

gtctogaaca ggagcgcgca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaacaggctc agcagggcca 3720

agcaggccgc gttctgaaa cacacgaagc agcagatcaa ggaatgcag ctttccttgt 3780

togatattgc gccgtggcgc gacacgatgc gagcgatgcc aaacgacaac gcccgctctg 3840
 ccctgttccac caccgcacaac aagaaaatcc cgcgcgaggc gctgczaaac aaggctcattt 3900
 tccacgtcaa caaggacgtg aagatcaacct acaccggcgt cgagctgcgg gccgaogatg 3960
 acgaactgggt gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020
 tcaccttcac gttctacgag ctttgccagg acctgggctg gtogatcaat ggccggtatt 4080
 acacgaaggc cgggaatatc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140
 accgcgttgg gcacctggaac togggtgcgc tgetgcaccg ctccgcylc ctggaccgtg 4200
 gcaagaaaac gtccogttgc caggctctga tcgacgagga aatgcctgtg ctgtttgctg 4260
 gcgaccacta cagcaaatc atatgggaga agtaccgcaa gctgtcgcgc acggcccgac 4320
 ggtatgttca ctatttcagc tcgcaccggg agccgtaccg gctcaagctg gaaaccttc 4380
 gcctcatgtg cggatcggtat tccaccgcgc tgaagaagtg gcgcgagcag gtgcgcgaag 4440
 cctgcgaaga gttgcgaggc agcggcctg tggaaacgc ctgggtcaat gatgacctg 4500
 tgcattgcaa acgctagggc cttgtgggt cagttccggc tgggggttca gcagccagcg 4560
 ctttactggc atttcaggaa caagcgggca ctgctcgac cacttgcctc gtcagatc 4620
 gctcgggacg caccgcgcgc tctaogact gccgataaac agaggattaa aattgacaat 4680
 tgtgattaa gctcaatcc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740
 cogattgtcg gccctgaaga aagctccaga gatgttcggg tcogtttacg agcacgagga 4800
 gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcgggccta 4860
 catcgacgcg gagatcattg ggctgtcggc ctccaacag gaggacggcc ccaaggacgc 4920
 tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gagggtcgc 4980
 cggtatgctg ctgcgggcgt tgcggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcatctt catctcggc gcacttaata ttctcgtatt 5100
ctggagcttg ttgtttatctt cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tctgtttcat ctctgcgct ctgctaggta gcccgatagc 5220
attgatggcg gtctctgggg ctatttgcgg aactgccccg gtggcgctgt tgggtgtgac 5280
accaaagcga gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga acogtgcga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgcttgccaa ctggcgccg gaggaactct gctcgttcca gtagctttag tgttlyalcc 5460
gccaatccg atgcctacag gaaccaatg tctcggcctg cgtggctcg gctgatcgg 5520
agcgggttta acctacttc ttgggttcg ggggatctcg cgactcgaa ctagagtgt 5580
ttccttactg ggctttctca gccocagatc tggggtcgat cagcgggga tgcacaggc 5640
cgacagtcgg aactccgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgct aacgttact tctaaagaaa tagcgcact cagcttctc agcggcttta 5760
tccagcgatt tctattatg tcggcatagt tctcaagatc gacagcctgt caggtttaag 5820
cgagaaatga ataagaagg tgataattcg gatctctcg agggagatga tatttgatca 5880
caggcagcaa cgtctgtca tegttaaat caacatgcta cctccgcga gatcatcgt 5940
gtttaaaco cggcagctta gttgcgctt tccgaatag catcggtaac atgagcaaa 6000
tctgcgctt tacaacggct ctcccgtga cgcgctccc gactgatggg ctgcctgtat 6060
cgagtgtga ttttgtcgg agctgccgtt cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgttt 6180
taatgtactg ggggtgtttt tcttttcacc agtgagacgg gcaacagctg attgccttc 6240
accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatgggtgt lcogaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtggt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagccccc atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaagga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgactcggg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtaccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtag gataagggcc agttccgccca 6900

gctcattaag agccagttca tgggogttgg catgatggcc gtcattgcac tgtaactcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgccttc gaatctgggt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagcttcc tcagccaggc ccagcccaag accgacaagg cctccctcca gaacgcagag 7140

aagaactgga ggggtgggtg caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgcaag aggatattct togactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatatcc tcaagtatct agagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaatata tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaagatttt cggcacggct acggaagacg gagaagccac cttcagtggg ctgcagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtggaga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactatata gaataagata ggtggagagc ttataccgag ctcccaaatc 7920

tgtccagatc atggttgacc ggtgcctgga tcttctata gaatcctct tatttugllya 7980

cctagctgat tctggagtga cccagagggt catgacttga gcctaaaatc cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatcgcg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccttc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcctcatcgg 8280

agaatatgga gcttcacga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagacg agtaccggc 8460

gcataaactc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgac cgggtgatga aaccggaaag gccgctcagg agctggccag cgggcgcagac 8580

cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640

tcagtccctg taggcagct ttgccccgtc tgtccgccg gtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtccaa catttgttgc catattttcc tgctctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttccatc caagaacctt 8820

tatttcccct aagtaagtac ttgtctacat ccatactcca tccttcccat cccattatcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgagc ttgactaaca gctaccocgc 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggogaagaat ctctgtcttt 9060

cagcttcgat gtagggaggg gtggatatgt cctgcgggta aatagotgag ccgatgggtt 9120

ctacaagat cgttatgttt atcggcactt tgcctcggcc gcgctccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctatlyc alctcccgcc gtgcacaggg 9240

tgctacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttgc gcccatcgg 9360

accgcaagga atcggatcat aactacatg gcgtgatttc atatgcgcga ttgtgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgctcc tcgcgcaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtcgggcacc tcgtgcacgc 9540

ggatttcgac tccaacaatg tctgacgga caatggcgcc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attccaata cgaggtcgcc aacatcttct tctggagggc 9660

gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcatctgt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

cgatccgga gccgggactg tcggggctac acaaatgcc gcgagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccc gcactcgtcc 9960

gagggcaaa gaatagagta gatgccgacc cggggatcga tccacttaac gttactgaaa 10020

tcacaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctogata agatacgttc atttgtccaa gcagcaaaaga 10140

gtgcotttota gtgattttaat agctccatgt caacaagaat aaaacgcggtt ttccgggttta 10200

octcttcag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagttaacgc 10260

cttncaggct cgggcaaga gaagaatagc ttagcagagc tattttcatt ttccgggagac 10320

ggatcaacgc agatcaacgc tegtcaagag acctacgaga ctgagggaac cgctcttggc 10380

tccacgcgac tatattttg tctotaattg tactttgaca tgcctctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtccuag cncctgggtt cgcaagata attgcatgtt 10500

tcttctctga actctcaacgc ctacaggaca cacattcacc taggtataa acctcgaaat 10560

canttcttac taagatggta tacaatagta accatgcacg ttgcctagt gaatgctccg 10620

taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg ttaccacaga 10680

atgcacaggt acacttggtt agaggtaac cttcttttota gctagaagtc ctogtgtact 10740

gtgtaagcgc ccactocaca tctccactcg acctgcaggc atgcaagctt aatctataca 10800

atgctccata gactcacatt gatattgtcg aagatttoga tgcctgactta gtagagcaac 10860

tacaaaagtt agcagagaag catgatttct taatctttga agacgcgag ttgcagata 10920

tgggtatgtg aattctatct attttttttc tgatgtgtgc atggatgact catgatcata 10980

ttcttaggta atactgtcaa gcacaaatat ggcaaggcgc ttacaagat tgcttcttgg 11040

tctcatatta ctaatgtcca cacagtctct ggagaaggta ttatcaaggg acttgccgaa 11100

tgcgctctcc ctcttggtcg tggcttgctt ttgctagcag aaatgtcacc tcaagggtca 11160

ttaaactaag gtatttaaac tgccgaatct gtcaaatatgg ctgcgcgcaa caaagatttc 11220

gtttttggct ttattgcaca acacaaaatg aatcagtatg atgatgagga ttttgttgtc 11280

atgtgcctg aagcttggcg taatcatggt catagctgtt tctgtgtga aattgttacc 11340

cgctcacaat tccacacaac atacgagcgc gaagcataaa gtgtaagcc tggggtgcct 11400
aatgagttag ctaactcaca ttaattgcgt tgcgctcact gcccgcttcc cagtcgggaa 11460
acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta 11520
ttggggcaaa gacaaaaggc cgacattcaa cgcattgagg gagggaaggc aaatattgac 11580
ggaaattatt cattaaaggc gaattatcac cgtaccgcac ttgagccatt tgggaattag 11640
agccagcaaa atcaccagta gcaccattac cattagcaag gccggaaaag tcaccaatga 11700
aaccatgatc agcagcaccg taalcagtag cgacagaatc aagtttgccct ttgagctcag 11760
actgtagcgc gttttctatg gcattttcgg tcatagcccc cttattagcg tttgccatct 11820
tttcataatc aaaatcacgc gaaccagagc caccacgcga accgcctccc tcagagccgc 11880
caccctcaga accgccaccc tcagagccac caccctcaga gccgccacca gaaccaccac 11940
cagagccgcgc gccagcattg acaggaggcc cgccttagta acatagatga caccgcgcgc 12000
gataatttat cctagtttgc gcgctatatt ttgttttcta tcgcgtatta aatgtataat 12060
tgcgggactc taatcataaa aaccatctc ataaataacg tcattgcatta catgttaatt 12120
attacatgct taacgtaatt caacagaaat tatatgataa tcacgcgaag accgycacac 12180
ggattcaatc ttaagaaact ttattgccaa atgtttgaac gatcggggat catccgggtc 12240
tgtggcggga actccagaa aatatccgaa cgcagcaaga tatcgcggtg catctcggtc 12300
ttgcctgggc agtcgccgcg gacgcggttg atgtggagcg cgggcccgat cataattgtg 12360
ctcaggatcg tggcgttgtg cttgtcggcc gttgctgtcg taatgatata ggcaccttcg 12420
accgcctgtt ccgcagagat cccgtgggcg aagaactcca gcattgagat ccgcgcctgg 12480
aggatcatcc agcggcggtc cgggaaaacg attccgaagc caaaccttcc atagaaggcg 12540
gggttgaatc cgaatctcgc tgatggcagg ttggggctcg cttggtcggt catttogaac 12600

cccagagtc cgcctcagaag aactcgtcaa gaaggcgata gaaggcgatg cgctgcgaat 12660
ggggagcggc gataccgttaa agcacgagga agcgggtcagc ccattcggcg ccaagctctt 12720
cagcaatatc acgggttagcc aacgctatgt cctgatagcg gtccgccaca ccagccggc 12780
cacagtcgat gaatccagaa aagcggccat ttccaccat gatattcggc aagcaggcat 12840
cgccatgggt cacgacgaga tcacgcgcgt cgggcgatcg cgccttgagc ctggcgaaca 12900
gttcgggtcg cgcgagcccc tgatgctctt cgtccagatc atcctgatcg acaagaccgg 12960
ctccatccg agtacgtgct cgtcogatgo gatgtttcgc ttggtggtcg aatgggcagc 13020
tagccggatc aagcgtatgc agccgccga ttgcacagc catgatggat accttctcgg 13080
caggagcaag gtgagatgac aggagatcct gccccggcac ttccccaat agcagccagt 13140
cccttccgc ttcagtaca acgtcgagca cagctgcgca aggaacgcc gtctggcca 13200
gccacgatag ccgcgctgcc tcgtcctgca gttcattcag ggcacgggac aggtcgggtc 13260
tgacaaaaag aaccgggcgc cctgcgctg acagccggaa cacggcgcca tcagagcagc 13320
cgattgtctg ttgtgccag tcatagccga atagcctctc caccacagcg gccggagaa 13380
ctcgcgtcaa tccatcttgc lcaatcatgc gaaacgatcc agatccggcg cagattatct 13440
ggattgagag tgaatatgag actctaattg gataccgagg ggaatttatg gaacgtcagt 13500
ggagcatctt tgacaagaaa tatttgctag ctgatagtga cctaggcgca cttttgaacg 13560
cgcaataatg gtttctgacg tatgtgotta gctcattaaa ctccagaaac ccgggctga 13620
gtggctcctt caacgttgcg gttctgtcag ttccaaacgt aaaacggctt gtcccgctc 13680
atcggcgggg gtcataacgt gactccotta attctcgcct catgatcaga ttgtcgtttc 13740
ccgccttcag tttaaactat cagtgtttga caggatatat tggcgggtaa acctaagaga 13800
aaagagcgtt tattagaata atcgatatt taaaagggcg lyaaggggtt tatccgttcg 13860

tccatttgta tgtgcatgcc aaccacaggg ttcccagat ctggcgccgg ccagcgagac 13920
gagcaagatt ggccgccgcc cgaacgacg cgacagcgcg ccagcacagc gtgcgcaggc 13980
aaattgcacc aacgcataca gcgccagcag aatgccatag tgggcggtga cgtcgttcga 14040
gtgaaccaga tcgcgcagga ggcccgccag caccggcata atcaggccga tgccgacagc 14100
gtcgagcgcg acagtgtca gaattacgat caggggtatg ttgggtttca cgtctggcct 14160
ccggaccagc ctccgctggc cagattgaac gcgcggatc tttatcactg ataagttggc 14220
ggacatatta tgtttatcag tgataaagtg tcaagcatga caaagttgca gcgaataca 14280
gtgatccgtg ccgccttgga cctgttgaa caggtcggcg tagacggtct gacgacacgc 14340
aaantggcgg aacgcttqcg ggttcagcag ccggcgcttt actggcactt caggaacaag 14400
cggcgctcgc tcgacgcact ggccgaagcc atgctggcgg agaataca gcattcggcg 14460
ccgagagccg acgacgactg gcgctcattt ctgacggga atgcgcgcag cttcaggcag 14520
gcgctgctcg cctaccgcga tggcgcgcg atccatgcg gcacgcgacc gggcgccacg 14580
cagatggaaa cggccgacgc gcagcttcgc ttctctcgcg agggcggttt ttccggccggg 14640
gacgcgctca atgcgctgat gacaatcagc tacttccactg ttggggccgt gcttgaggag 14700
caggccggcg acagcagtcg cggcgagcgc ggccggcacg ttgaacagcg tcgcctctcg 14760
ccgctgttgc gggccgcgnt agacgccttc gacgaagcgg gtccggacgc agcgttcgag 14820
cagggactcg cggtgattgt cgatggattg gcgaaaagga ggctcgttgt caggaaagtt 14880
gaaggaccga gaaagggtga cgattgatca ggcacgctgc cggagcgcaa cccactcact 14940
acagcagagc catgtagaca acatccctc cccctttcca ccgcgtcaga cgcccgtagc 15000
agcccgctac gggctttttc atgccctgcc ctacgctcca agcctcacgg ccgcgctcgg 15060
cctctctggc ggccctctcg cgtctctcgc ctctctcgt cactgactcg ctgcgctcgg 15120

togtttcggt cggcgagcg gtatcagctc actcaaaggc ggtaatacgg tttatccacag 15180
 aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccagggaacc 15240
 gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca 15300
 aaaatcgacg ctcaagtacg aggtggcgaa acccgacagg actataaaga taccaggcgt 15360
 ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgcgcgtt accggatacc 15420
 tgtccgcctt tctcccttcg ggaagcgtgg cgtttttccg ctgcataacc ctgcttcggg 15480
 gtcattatag cgatttttcc ggtatatcca tcctttttcg cagcatatac aggattttgc 15540
 caaagggttc gtgtagactt tccttgggtg atccaacggc gtcagccggg caggataggt 15600
 gaagtagacc caccgcgag cgggtgttcc ttcttcaactg tcccttatcc gcacctggcg 15660
 gtgctcaacy ggaatcctgc tctgcgagcg tggcgggcta ccgcgcgcgt aacagatgag 15720
 ggcaagcgga tggctgatga aaccaagcca accaggaagg gcagcccacc tatcaagggtg 15780
 tactgccttc cagacgaacg aagagcgatt gaggaaaagg cggcgggcgg cggcatgagc 15840
 ctgtcggcct acctgctggc cgtcggccag ggctacaaaa tcacgggcgt cgtggactat 15900
 gagcacgtcc gcgagctggc ccgcatcaat ggagacctgg gccgcctggg cggcctgctg 15960
 aaactctggc tcaccgacga ccgcgcacg gcgcgggttg gtgatgccac gatcctcgcc 16020
 ctgctggcga agatcgaaga aaacaggac gagcttgcca aggtcatgat gggcgtggtc 16080
 cgcgcagggg cagagccatg acttttttag ccgctaaaaa gcccgggggg tgcgcgtgat 16140
 tgccaagcac gtccccatgc gctccatcaa gaagagcgac ttgcgggagc tggtagagta 16200
 catcaccgac gagcaaggca agaccgagcg cctttgcgac gctca 16245

<212> DNA

<213> Artificial

<220>

<223> Promotor

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 37

ccgggctggt tgcctcgcc gctgggctgg cggcgtcta tggccctgca aacgcgcag	60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgcg gogttgtgga tacctcgcg	120
aaaacttgcc cctcactgac agatgagggg cggacgttga caactgaggg gcgcactcac	180
ccggcgcgcc gttgacagat gaggggcagg ctcgatttcg gcggcgagc tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag ttccccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gagggcgcg atccttgaca cttgaggggc agagtgtgta cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaaggtt ttccgccct	480
ttttcgccca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcggt gaccgcgcac gcgaagggg ggtgccccc	600

ctttctogaac cctcccggcc cgctaacgcg ggctcccat cccccaggg gctgcccc 660

tggcgcgga acggcctcac cccaaaaatg gcagcgctgg cagtccttgc cattgcggg 720

atcggggcag taacgggatg ggcgatcagc cggagcgga cggccggaag cattgacgtg 780

ccgcagggtg tggcatcgac attcagcagc caggtgccgc ccaatgagga cggcgccgtg 840

gggtggcgcc tgccttcac ttcggcgtc ggggcattca cggacttcat gggggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtccgg gtgccgtgt cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaa 1020

acgagaattg gacctttaca gaattactct atgaagcgc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatatgac aatactgata 1140

agataatata tcttttatat agaagatgc gcgtatgta aggatttcag ggggcaaggc 1200

ataggcagcy cgtttatcaa tatatctata gaatgggcaa agcataaaaa ntgcattgga 1260

ctaattgctt aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttcgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgcgctc aattcgctgc gtatatcgt tgcgtattac gtgcagcttt 1500

cccttcaggc gggattcata cagcgccgag ccacccgca tccatatcac cactcaag 1560

ggtgacagca ggtcataag acgcccagc gtcgccatg tgcgttcacc gaatacgtgc 1620

gcaacaacgg tottccggag actgtcatc gcgtaaaaca gccagcgctg gogogattta 1680

gccccgacat agcccactg ttctgcatt tccgcgcaga cgatgacgc actgcccggc 1740

tgtatgcgcg aggttaaccg ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcatacaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgtgat gtccggcgggt gcttttgcg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100

aaatcggtc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta tttaagttt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220

aattagcttc ttgggtatc ttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcacog aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaaggata taagctggg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacacgc ggtataaagg gaaccactat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctccc ttttccaaag gtccctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcccttgcct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgcg tcaggtctct 2640

tcactccatc gacatatacg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agccccga 2820

ggaacttgct ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgato ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctgcgtc cgttcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tatttatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
 tcttcgcgat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
 ggtcgctggt attcgtgcag ggcaagatc ggaataccaa gtacgagaag gacgcccaga 3240
 cggctctcgg gaccgacttc attgccgata aggtggatta tctggacac aaggcaccag 3300
 gcgggtcaca taaggataa gggnatattg cccgcgcgtg agtcggggca atccgcgaag 3360
 gagggtgaat gaatcggaag tttagccgga aggcatacag gcaagaactg atcgagcgg 3420
 ggttttcgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480
 aaaccttcca gtccgtggc tcgatggtc agcaagctac ggccaagatc gagcgcgaca 3540
 gcgtgcaact ggtcccccgt gccctgcccg cgcctcggc gcgctggag cgttccgctc 3600
 gtctcgaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaaacta 3660
 tgacgaccaa gaagcgaaaa accgcggcg aggaacctgg aaaacagtc agcgaggcca 3720
 agcaggccgc gttgctgaaa caacaagaag agcagatcaa ggaatgcag ctttccttgt 3780
 tcgatattgc gcgctggcg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840
 cctgttccac cagcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900
 tccagctcaa caaggacgtg aagatcacct acaccggcgt cgagctggcg gccgacgatg 3960
 acgaactggt gtygcagcag gtgttgagc acggaagcg caccctatc gccgagccga 4020
 tcaaccttcc gttctacgag ctttgccagg acctgggctg gtcgatcaat gcccggtatt 4080
 acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac gccgatggcg ttcacgtccg 4140
 accgcgttgg gcacctggaa tcygtgcgc tgcgtgcaac ctcccgctc ctggaccgtg 4200
 gcaagaaaaac gtcccgttgc caggtcctga tcgacgagga aatcgctgctg ctgtttgctg 4260
 gcgaccacta caggaattc atatgggaga agtacgcaaa gctgtcgccg acggcccgcac 4320

ggatgttcga ctatttcago tcgcacccgg agccgtaccc gctcaagctg gaaaccttc 4380
 gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
 cctcggaaga gttgcgaggo agcggcctgg tggaaacacgc ctgggtcaat gatgacctgg 4500
 tgcattgcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560

 ctttactggc atttcaggaa caagcgggca ctgctcgacg caettgcttc gtcagtatc 4620
 gctcgggaag cagcgcgcg tctacgaact gcgataaac agaggattaa aattgacaat 4680
 tgtgattaag gctcagattc gacggcctgg agcggccgac gtgcaggaLL lccgcgagat 4740
 ccgattgtog gccctgaaga aagctccaga gatgttcggg tccgtttacg agcagcagga 4800
 gaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tggcgccta 4860
 catcgacgcg gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920
 tcacaaggcg catctgtccg gcgttttctg ggagcccgaa cagcgaggcc gaggggtcgc 4980
 cggtatgctg ctgcgggcgt tgcggcggg tttattgttc gtgatgatcg tccgacagat 5040
 tccaacggga atctgggtga tgcgcattct catctcggc gcacttaata tttcgtatt 5100
 ctggagcttg ttgtttattt cggctacccg cctgcggggc ggggtcggcg cgacggtagg 5160
 cgtgtgtcag ccgctgatgg tcgtgttcac ctctgccgt ctgctaggta gcccgatacg 5220
 attgatqccg gtcctggggg ctatttcggg aactgcgggc gtggcgctgt tgggttgac 5280
 accaaacgca cgcctagatc ctgtcggcgt cgcgcgggc ctggcggggg cggtttccat 5340
 ggctgtcgga accgtgtcta ccgcgaagtg gcaacctccg gtgcctctgc tcaccttac 5400
 cgcttgga caaggcggcg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
 gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatccg 5520
 agcgggttta acctacttcc ttggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggcttttctca gcccagatc tggggtcgat cagccgggga tgcctcaggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttctc agcggttcta 5760

tccagcgatt tctattatg tcggcatagt tctcaagatc gacagcctgt cagcggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgog agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacat caacatgcta cctccgcga gatcatccgt 5940

gtttcaaac ccgagcctta gttgccgttc ttcgaatag catcggtaac atgagcaaa 6000

tctgcgcct tacaacggct ctcccgctga cgcgcctccg gactgatggg ctgcctgtat 6060

cgagtggtag ttttctgcc agctgccggg cggggagctg ttggctggct ggtggcagga 6120

tatatgtgg tgtaacaaa ttgacgctta gacaacttaa taacacattg cggagctttt 6180

taetgtactg ggggtggttt tottttcacc agtgagacgg goaacagctg attgcccttc 6240

accgcctgac cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatgggtgt tccgaaatcg gcaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aayaacgttg 6420

actccaacgt caaaggcgga aaaacgtct atcagggcga tggccacta cgtgaaccat 6480

cacccaaatc aagttttttt gggtcgaggt gccgtaaac actaaatcgg aacctaaag 6540

ggagccccgc atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgaact gttgggaagg gcgatcggtg 6660

cgggcctctt cgctattacg coagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cggtgtaaaa cgaaggccag tgaattcgag 6780

ctcggtaccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgccca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgccttc gaactcgggt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagtcctt attgaagtcg gaggacggag 7200

cgggtgcaag aggatattct tcgaclclyl attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttoca ggctgagact ctagcttggg gcatagaggg 7320

tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttgaacgtg 7500

gcactattga tcattccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tccttagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggtc acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgtt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtggaga agtggaacta aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaatc 7920

tgtccagatc atggttgacc ggtgcctgga tcttctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggc catgacttga gcctaaaatc cgcgcctcc 8040

accatttga gaaaaatgtg acgaactcgt gagctctgta cagtgcacgg tgactctttc 8100

tgccatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggtcttgag gtgcagtgga tgattattaa tcggggaccg gcccccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg 8280

agaatatgga gcttcacga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaata gcatgccatt aacctaggta cagaagtcca attgcttcgg 8400

atctggtaaa agattcacga gatagtacct totcgaagt aggtagagcg agtaccggc 8460

gcgtaagctc cctaattggc ccatacggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgagac 8580

cgggaaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640

tcagtcacctg gtaggcagct ttgcccctgc tgtccgcccg gtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtcaca cttttgttc catattttcc tgcctctccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttccctc aagtaagtac ttgtctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccatc tcatcgagc ttgactaaca gctaccocgc 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatga 9000

aaagttoagc agcgtctccg acctgatgca gctctcggag ggogaagaat ctctgtgttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggaactt tgcctgggc gcgctccoga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgcct gttctgcagc cggctgcgga 9300

ggccatggat ggcagcgtg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtaaat acactacatg gcgtgatttc atagcgcgga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tccgcagggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgcccga gtcggcacc tctgacacgc 9540

ggatttcggc tccaacaatg tctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attccaata cgggtcgc aacatcttct tctgggggc 9660

gtggttggt tgtatggagc agcagacgc ctacttcgag cggaggcacc cggagcttgc 9720

aggatgcgc cggctccggc cgtatatgct cgcattggt cttgaccaac tctatcagag 9780

cttggttgac gccaatttc atgatgcgc ttggcgcgag ggtcgatgc acgcaatcgt 9840

ccgatccgga gccgggactg tggggcgtag acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gcaactgtcc 9960

gagggcaag gaatagagta gatccgacc ggggatcga tccactaac gttactgaaa 10020

tcatcaaaa gcttgacgaa tctggatata agatcgttg tgcgatgtc agctccggag 10080

ttgagacaaa tgggtttcag gatctcgata agatacgttc atttgtccaa gcagcaaga 10140

gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttcagc atacagctca tctgcaatgc attaatgc atgactgcaac ctagttaacg 10260

cttcnaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgtctcttgc 10380

tccagcgac tatatatattg tctctaattc tactttgaca tgcctctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaagata attgcatgtt 10500

tcttcttga actctcaagc ctacaggaca cacattcctc gtaggtataa acctcgaaat 10560

cantctctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctcgg 10620

taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680

atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctctgtgtact 10740

gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt ttttcgagtt 10800

tttttttttt ttctttgtga aggatttatt gttattggta tccatttttt attggaagac 10860

aagataagtt aatattgatt ttgcttaag attaaaagga aatcagaaaa cgacaataaa 10920

aatgtaacg gacaaactat ggtgtcgatt ataagtctaa atccttaaaa aatgacaacg 10980

agttgcttc ctctgaaaaa'aattcttttg tctttgcaag aaaggtttct tttttgttg 11040

cttgcatcac ttaaacatca aatcaaatga aaggaataaa gcagatttga gggcgaataa 11100

ggattttctg gtaacaaga tgtgagtgac acctaaaggaa ctaaattgcca ttcatttggt 11160

ttaaaaagac atcaagatt gatgatcaac aggattgaga gagagaaaaa gaactcgtgt 11220

catttttttc tgttgactga aattttatat ttagaaaaaa tgtcaaatct atagctttag 11280

ctatattaca taacatttga aataataata ataaaaaaag acacattaga gacacttttc 11340

aaactctaaa taactgtcta taaacacaaa gaaaacaaag acctctataa caacttatta 11400

gattttttct gtacttttgt ctaaagatga tgtattcttg ttatccaca cttctttcat 11460

ttgtttctga tgctactaaa tatacaaat ttcttttttg caagagatat tattcaaaa 11520

attttcaaaa agaaattttt ttcaaatag cagttgatcg tgtaaccaa agaggttctt 11580

tggtattttg cacttcgcgt ttgcggtgat gcatattcaa agtaatatat ggaataaaca 11640

acgtgtttaa gcatgaaaga aaggaaacaa aggcgcgttt gaacaaatgc ataattttc 11700

agacaaaaat gatctaagc aagcagtaaa tcaacaaga aacattgctg attcgggtta 11760

gaaacagata aaagtctaata aagccactaa gtatacttca atgaactttt tgtatgctta 11820

tggtccaatc agaccaataa ttgtgacca ttcttgaggt ggcttttggtg atgcggaaac 11880
agaaaaaaat tttctcacca atcgatttaa aaaacaattt ctgctttgaa ccaaaacttt 11940
ttttttctct ttaatcatta accttatcaa gtatgtacct acctcaaaag tctcactca 12000
agcacaatta tgctaacatt gtccacott ctctttagaa atgtgtcga agctgcagtc 12060
aatcaacgtc aagcccgccg cgtttgaact agcccgcgac atcaacgggc ccaagtcgtg 12120
cctgcatgct cagcgggtgt cgttagttcg gctgcgagtg gcagcaccac agacagagga 12180
ggcgctggga accgtgcagg ctgcggcgcc gggcgatgag cacagcgccg atgtagcact 12240

ccagcagctt gacggggcta togcagagcg tctgtcccg cgcaaacggg agcagctgtc 12300
ataccagggt gcgcgcattg cagcatcaat tggcgtgtca ggcattgcca tcttcgcac 12360
ctacctgaga ttgtcatgc acatgacctt gggcgccgca gtgcgatggg gtgaagtggc 12420
tggcactctc ctcttgggtg ttggtggcg gctcggcatg gagatgtatg ccgctatgc 12480
acacaaagcc atctggcatg agtcgcctct gggctggctg ctgcacaaga gccacacac 12540
acctcgactt ggacctttt aagccaacga ctgttttga atcatcaatg gactgcgcg 12600
catgctcctg tgtaccttg gtctctggt gcccaacgtc ctggggggcg cclgulllg 12660
agcggggctg ggcacacgc tatacggcat ggcataatg tttgtacag atggcctggt 12720
gcacaggcgc tttccaccg ggcacatgc tggcctgcc tacatgaagc gctgcacagt 12780
ggccaccag ctacaccaca gcggcaagta cgggtggcgcc cctggggta tgttcttggg 12840
tccacaggag ctgcagcaca ttccaggtgc ggcggaggag gtggagcgac tggctctgga 12900
actggacttg tccaagcgtt agaagcttgg cgtaatcatg gtcatactgt tttctgtgt 12960
gaaattgtta tccgctcaca attccacaca acatacgagc cgggaagcata aagtgtaaag 13020
cctggggctg ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgctt 13080

tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag 13140
gcgggtttgcg tattggggcca aagacaaaag ggcgacattc aaccgattga gggagggaag 13200
gtaaatattg acggazatta ttcattaaag gtgaattatc accgtcacgc acttgagcca 13260
tttgggaatt agagccagca aaatcaccag tagcaccatt accattagca aggcgcgaaa 13320
cgtcaccaat gaaaccatcg atagcagcac cgtaatcagt agcgacagaa tcaagtttgc 13380
ctttagcgtc agactgtagc gcgttttcat cggcattttc ggtcatagcc cccttattag 13440
cgttttgcat cttttcataa tcaaaatcac cygaaccaga gccaccacgc gaaccgcctc 13500
cctcagagcc gccaccctca gaaccgccac cctcagagcc accaccctca gagccgccac 13560
cagaaccacc accagagccg ccgcacgatc tgacaggagg ccgatctag taacatagat 13620
gacaccgcgc gcgataatct atcctagttt gcgcgctata ttttgttttc tatcgcgtat 13680
taaatgtata attgcgggac tctaataata aaaaccatc tcaataataa cgtcatgcat 13740
tacatgttaa ttattacatg cttaacgtaa ttcaacagaa attatatgat aatcatcgca 13800
agaccggcaa caggattcaa tottaagaaa ctttattgcc aaatgtttga acgatcgggg 13860
atcatccggg tctgtggcgg gaactccacg aaaatatccg aacgcagcaa gatatcggcg 13920
tgcatctcgg tcttgctcgg gcagtcgcgc ccgacgcgct tgatgtggac gccggggccc 13980
atcatattgt cgctcaggat cgtggcggtg tgcttgctcg ccgttgctgt cgtaatgata 14040
tcggacacct cgaccgcctg ttccgcagag atcccgtggg cgaagaactc cagcatgaga 14100
tccccgcgct ggaggatcat ccagccggcg tcccggaata cgattccgaa gcccaacctt 14160
tcatagaagg cggcggttga atcgaaatct cgtgatggca ggttgggcgt cgcttggtcg 14220
gtcatcttga accccagagt ccgcctcaga agaactcgtc aagaaggcga tagaaggcga 14280
tgcgctgcga atcggggagc gcgataccgt aaagcacgag gaagcggtea gccatttcgc 14340

cgccaagctc ttcagcaata tcacgggtag ccaacgctat gtcctgatag cggctccgca 14400
caccagccg gccacagtcg atgaatccag aaaagcggcc attttccacc atgatattcg 14460
gcaagcagggc atcgccatgg gtcacgacga gatcatcgcc gtcgggcatg cgcgccttga 14520
gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc ttgtccaga tcactcctgat 14580
cgacaagacc ggcttccato cagtaoctg ctcgctcgat gcgatgttcc gcttggtggt 14640
cgaatgggca ggtagccgga tcaagcgtat gcagccgcg cattgcacga gccatgatgg 14700
atactttctc gccaggagca aggtgagatg acaggagatc ctgccccggc acttgcacca 14760
atagcagcca gtcccttccc gcttcagtg caacgtcgag cacagctcgc caaggaaagc 14820
ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg cagttcattc agggcaccgg 14880
acaggtcggc cttgacaaaa agaaccgggc gccctgcgc tgacagccgg aacacggcgg 14940
catcagagca gccgattgtc tgttggtccc agtcatagcc gaatagctc tccaccaag 15000
cgcccgaga acctgcgtgc aatccatctt gttcaatcat gcgaaacgat ccagatccgg 15060
tgcagattat ttggattgag agtgaatatg agactctaat tggataccga ggggaattta 15120
tggaacgtca gtggagcatt ttgacacga aatatttgct agctgatagt yaccttaggc 15180
gacttttgaa cgcgcaataa tggttttctga cgtatgtgct tagctcatta aactccagaa 15240
accgcgggct gagtggctcc ttcaacgttg cggttctgtc agttccaaac gtaaacgggc 15300
ttgtcccgcg tcacgcggcg gggtcataac gtgactccct taattctcgc ctcgatgaca 15360
gattgtogtt tccgccttc agtttaaaact atcagtggtt gacaggatat attggcgggt 15420
aaacctaga gaaaagagcg ttattagaa taatcggata tttaaaaggg cgtgaaaagg 15480
tttatcgtt cgtccatttg tatgtgcag ccaaccacag ggttccccag atetggcgcc 15540
ggccagcgcg acgagcaga ttggccgccc ccgaaacga tccgacagcg cggccagcac 15600

aggtagcgag gcaaattgca ccaacgcata cagcgccagc agaatgccat agtgggagggt 15660

gacgtcgttc gagtgaacca gatcgcgag gagggccggc agcaccggca taatcaggcc 15720

gatgcgcaga gcgtcgagcg cgacagtgtc cagaattacg atcaggggta tgttgggttt 15780

cacgtctggc ctccggacca gcctccgctg gtcgattga acgcgcgat tctttatcac 15840

tgataagttg gtggacatat tatgtttatc agtgataaag tgtcaagcat gacaagttg 15900

cagccgaata cagtgatccg tgcgccctg gacctgttga acgaggtcgg cgtagacggc 15960

ctgacgacac gcaaaclygc ggaacgggtg ggggttcagc agcggcgct ttaactgcac 16020

ttcaggaaca agcgggagct gtcgacgca ctggccgaag ccattgctggc ggaagaatcat 16080

acgcattcgg tgcgagagc cgacgacgac tggcgctcat ttctgatcgg gaatgccgcg 16140

agcttcaggc agggcgctgt cgccctaccg gatggcgcg gcattccatgc cggcacgcga 16200

cggggcgac cgcagatgga aacggccgac gcgcagcttc gcttcctctg cgaggcgggt 16260

ttttggcgg gggagccgt caatgcgctg atgacaatca gctacttcac tgttggggcc 16320

gtgcttgagg agcaggcgg cgacagcgat gccggcgagc gggggcgac cgttgaacag 16380

gtccgctct cgccgctgtt gccggcgcg atagacgctc tcgacgaagc cggtcgggac 16440

gcagcgttcg agcagggact cgcggtgatt gtcgatggat tggcgaaaag gaggctcgtt 16500

gtcaggaacg ttgaaggacc gagaaggggt gacgattgat caggacggct gccggagcgc 16560

aaccactca ctacagcaga gccatgtaga caacatccc tcccccttc caccggtca 16620

gaagcccgta gaagccgct aggggtttt tcatgccctg cctagcgtc caagctcac 16680

ggcgcgctc ggctctctg gcggcttct ggctctctc cgttcctcg ctactgact 16740

cgctcgctc ggtggttcgg ctggcgag cggtatcagc tcaactcaag gcggtaatac 16800

ggttatccac agaatacagg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa 16860
aggccaggaa ccgtaaaaaa gccgcgttgc tggcggtttt ccataggctc cgccccctg 16920
acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa 16980
gataccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctggcgc 17040
ttaccggata cctgtccgcg tttctccctt cgggaagcgt ggcgcttttc cgtgcataa 17100
ccctgcttcg gggtcattat agcgattttt tcggtatata catccttttt cgcacgatat 17160
acaggatttt gccaaagggc tcgtgtagac tttccttggt gtatccaacg gcgtcagccg 17220
ggcaggatag gtgaagtagg cccacccgcg agcgggtgtt ccttcttcac tgtcccttat 17280
tcgcacctgg cgggtctcaa cgggaatcct gctctgcgag gctggccggc taacgccggc 17340
gtaacagatg agggcaagcg gatggctgat gaaaccaagc caacaggaa ggcagccca 17400
cctatcaagg tgtactgcct tccagacgaa cgaagagcga ttgaggaaaa ggcggcgggc 17460
gccggcatga gctgtcggc ctactgctg gcgctcgcc agggctacaa aatcacgggc 17520
gtcgtggact atgagcacgt ccgcgagctg gccgcacatc atggcgacct gggccgcctg 17580
ggcggcctcg tgaaactctg gctcacgcac gaccgcgcga cggcgcggtt cggatgatcc 17640
acgatcctcg ccctgctggc gaagatcgaa gagaagcagg acgagcttgg caaggtcatg 17700
atggcgctgg tccgcccagg ggcagagcca tgactttttt agcgctaaa acggcgggg 17760
ggcgcgctg attgccaac acgtcccat gcgctccatc aagaagagcg acttcggga 17820
gctggtgaag tacatcacg acgagcaagg caagaccgag cgccttttgc acgctca 17877

<210> 38

<211> 17238

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 38

```
ccgggctgggt tgcctctgcc gctgggctgg cggcctgcta tggcctcgca aacgcgccag      60
aaacgcgctc gaagcgtgtt gcagacacc ccagccgcgc gcgttgtgga tacctcgccg      120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac      180
cggcgccggc gttgacagat gaggggcagg ctgcatttcg gccggcgacg tggagctggc      240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag ttcccacag atgatgtgga      300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat      360
gagggcgccg atccttgaca cttgaggggc agagtgtgta cagatgaggg gcgcacctat      420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccg      480
ttttcgccca ccgctaaact gtcttttaac ctgcttttaa accaatattt ataaaccttg      540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc      600
ctttctgaac cctcccgccc cgctaacgcg ggcctcccat cccccaggg gctgcgcgcc      660
```

tgggccgcga acggcctcac cccaaaaatg gcagcgctgg cagtccctgc cattgccggg 720

atcggggcag taacgggatg ggcgatcagc ccgagcgcca cggccggaag cattgacgtg 780

ccgcaggctg tggcatcgac attcagcgac cagggtccgg gcagtgaagg cggcgccctg 840

ggtggcgccc tgcccctcac ttccggcgtc ggggcattca cggacttcat gggggggccg 900

gcaattttta ccttgggcac tcttgcata gtggtccgg gtgccgtgct cgtgttcggg 960

ggtgogataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgc alatttataa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatata gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgtttatcaa tatactata gaatgggcaa agcataaaaa cttgcattga 1260

ctaattgctg aaaccaggga caataacctt atagcttga aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac ttgtcatgac 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgcgcctc aattcgctgc gtatatcgct tgctgattac gtgcagcctt 1500

cccttcaggc gggattcata cagcgccag ccactcgcca tccatatcac cagtcgaaag 1560

ggtgacagca ggctcataag acgcccagc gtgcctatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttcggag actgtcatc gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttccgtccatt tccgcgcaga cgtgacgtc actgccccgc 1740

tgtatgcgcg aggttaccca ctggggcctg agttttttta gtgacgtaaa atcgtgttga 1800

ggcaacgcc cataatgcgg gctgttgccc ggcatacaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtacgg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgctgat gtcggcggt gcttttgccg ttaacgacca 1980
 ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
 aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtgttttca 2100
 aaatcggtc cgctgatact atgtttatcg ccaactttga aaacaacttt gaaaagctg 2160
 ttttctggta ttttaagttt tagaatgcaa ggaacagtga attggagtgc gtcttgttat 2220
 aattagcttc ttgggttato tttaaatact gtgaaaaga ggaaggaaat aataaatggc 2280
 taaatgaga atatcaccgg aattgaaaa actgalyaa aaataccgct gcgtaaacga 2340
 tacggaagga atgtctctcg ctaaggata taagctggtg ggagaaaatg aaaacctata 2400
 tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
 catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520
 tgatggctgg agcaatctgc tcattgagta ggccgatggc gtcccttgct cggagagta 2580
 tgaagatgaa caagccctg aaaagattat cgagctgtat cgggagtga tcaggctctt 2640
 tcactccatc gacatatacg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
 attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggagaagaa 2760
 cactccattt aaagatccgc cggagctgta tgatttttta aagacggaaa agcccgaaga 2820
 ggaacttgto ttttccacg gcgacctggg agacagcaac atcttttgta aagatggcaa 2880
 agtaagtggc ttattatgc ttgggagaag cggcagggcg gacaagtggc atgacattgc 2940
 cttctcgctc cggctgalca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000
 tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060
 attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
 tcttcgcat caagtgtttt ggcctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggctcgctggt attcgtgcag ggcaagattc ggaataocaa gtacgagaag gacggccaga 3240
 cgggtctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
 gcgggtcaaa tcaggaataa gggcacattg ccccgcgctg agtcggggca atcccgoaag 3360
 gagggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
 ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgctcatgctg gcgccccgcg 3480
 aacccctcca gtccgtcggc togatggctc agcaagctac ggccaagatc gagcgcgaca 3540
 gcgtgcaact ggctccccc gccctgucuc cyccatcgcc cgcctggagc cgttcgctgc 3600
 gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaaacta 3660
 tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720
 agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaatgacg ctttccctgt 3780
 tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840
 cccgtgtcac cagcgcaac aagaaaatcc cgcgcgagggc gctgcataac aaggtcattt 3900
 tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgaactgcgg gccgacgatg 3960
 acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagucga 4020
 tcacottcac gttctacgag ctttgccagg acctgggctg gtccgatcaat ggccggatatt 4080
 acacgaaggc cgaggaaatgc ctgtcgccgc tacaggcgac ggcgatggcg ttcacgtccg 4140
 accgcgttgg gcacctggaa tcgggtgcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200
 gcaagaaaaac gtcccgttgc caggtcctga tcgacgagga aatcgctctg ctgtttgctg 4260
 gcgaccacta cacgaaatcc atatgggaga agtacccgaa gctgtgcgcc acggcccgac 4320
 ggatgttcga ctatttcagc tcgcacgggg agccgtaccc gctcaagctg gaaaccttc 4380
 gcctcatgtg cggatcggat tccaccggcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaaacacgc ctgggtcaat gatgacctgg 4500
 tgcattgcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560
 ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttcttcg gtcagttatc 4620
 gctcgggacg caccgcgcg cctacgaact gccgataaac agaggattaa aattgacaat 4680
 tgtgattaag gtcagattc gacggcttgg agcggcggc gtgcaggatt tccgcgagat 4740
 ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcagagga 4800

 gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tggcgcccta 4860
 catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920
 tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980
 cggtatctcg ctgcggcggt tgcggcgggg tttattgtct gtgatgatcg tccgacagat 5040
 tccaaaggga atctggtgga tgcgcattct catcctcgcc qcacttaata tttcgtatt 5100
 ctggagcttg ttgtttattt cggctacacg cctgcggggc ggggtgcggg cgacggtagg 5160
 cgctgtgcag ccgctgatgg togtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
 attgatggcg gtccctggggg ctatttgcgg aactgcgggc gtggcgctgt tgggtgtgac 5280
 accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
 ggcgctcgga accgtgctga ccgcgaagtg gcaacctccc gtgcctctgc tcacctttac 5400
 cgccgggcaa ctggcggcgg gaggacttct gctogttcca gtagctttag tgtttgatcc 5460
 gccaatcccc atgcctaacg gaadcaatgt tctcggcctg gcgtggctcg gctgatacgg 5520
 agcgggttta acctacttcc tttggttccg ggggatctcg cgaactogaac ctacagttgt 5580
 ttccctactg ggctttctca gccccagatc tggggtcgat cagccgggga tgcacagggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcaact tctaaagaaa tagcgccact cagcttcctc agcggtctta 5760

tcacgcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cagcggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta cctccgcga gatcatccgt 5940

gtttcaaac ccgagcgtta gttgccgttc ttccgaatag catcggtaac atgagcaaaag 6000

tctgcgcct tacaacgggt ctcccgctga cgcgctccg gactgatggg ctgcctgtat 6060

cgagtgggtg ttttgtgcc agctgccggt cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaaaaa ttgacgtta gacaactta taacacattg cggacgtttt 6180

taatgtactg ggggtggttt tcttttoacc agtgagacgg gcaacagctg attgccttc 6240

accgcctggc cctgagagag ttgcagcaag cggttccacg tggtttggcc cagcaggcga 6300

aatcctgtt tgatggtgt tccaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

cogagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaaggcgga aaaaccgtct atcaggggcga tggccacta cgtgaaccat 6480

cacccaatc aagttttttt gggtcgaggt ggcgtaaagc actaaatcgg aaccctaag 6540

ggagccccc atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgacgtgtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaa gcgattaagt 6720

tgggtaaogo cagggttttc cagtcacga cgttgttaaa cagggcgag tgaattcag 6780

ctcggatccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactacca catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcggttg catgatggcc gtcatgcate tgtaactcaa 6960

gtacaccaac gctcttctga tccagtgat catccgtga aggcgcttc gaatctggtt 7020

aagatccaag tcttcgggaa gcaagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagcttct tcagccaggg ccagcccaag accgacaagg cctccctcca gaaagcgag 7140

aaagaactgga cccctcgtgt caaggaggag taagctcctt attgaagtgc gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctgcttgga gcatagaggg 7320

tcctttggct ttcaatatcc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatgc catgaggttt 7440

tcgaaatata tccggatgto gaaggcttgg ggcacctgog ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaa ggcggttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggttggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctogagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactatata gaataagata ggtggagagc ttatacogag ctcccaaatc 7920

tgtccagato atgggtgaac ggtgcctgga tcttctata gaatcatcct tattogttga 7980

cctagctgat tctggagtga cccagagggt catgacttga gctaaaato cccgcctcc 8040

accatttcta gaaaaatgtg acgaactcgt gagctctgta cagtaccggg tgactcttct 8100

tggcatcgcg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctctgg cggctctcgag gtgcagtgga tgattattaa tccgggaccg gcgcccctc 8220
 cgccccgaag tggaaaaggct ggtgtgcccc tctgtgacca agaattctatt gcatcatcgg 8280
 agaatatgga gcttcatcga atcacoggca gtaagcgaag gagaatgtga agccagggggt 8340
 gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttcgg 8400
 atctgtgtaa agattcacga gataglacct tctcogaagt aggtagagcg agtaccggcg 8460
 gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
 ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgccagc 8580
 cgggaacaca agctggcagt cgaccatcc ggtgctctgc actcgacctg ctgaggtccc 8640
 tcagtcctcg gtaggcagct ttgcccgtc tgtccggcg gtgtgtcggc ggggttgaca 8700
 aggtcgttgc gtcagtcaca catttgttgc catattttcc tgcctctccc accagctgct 8760
 cttttctttt ctctttttt tcccatcttc agtatattca tcttccatc caagaacctt 8820
 tatttccct aagtaagtac ttgtctacat ccatactcca tcttcccat ccttattcc 8880
 tttgaacctt tcagttcgag ctctccact tcctgcgcgc ttgactaaca gctaccgcgc 8940
 ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaacl tctgatcga 9000
 aaagtctgac agcgtctcgc acctgatgca gctctcggag ggcgaagaat ctggtgcttt 9060
 cagcttcgat gtaggggggc gtggatatgt cctgcgggta aatagctcgc ccgatggttt 9120
 ctacaaagat cgttatgttt atggcactt tgcatcgcc gcgctccga ttcggaagt 9180
 gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240
 tgtcacgttg caagacctgc ctgaaaccga actgcccgt gttctgcgc cggtcgcgga 9300
 ggcacatggat gcgctcgtg cggcgatct tagccagacg agcgggttcg gccattcgg 9360
 accgcaagga atcgggtcaat acactacatg gcgtgatttc atatgcgoga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacacogtc agtgcgtccg tgcgcagggc 9480

tctcgtatgag ctgatgcttt gggccgagga ctgcccogaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcttgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attccaata cgaggctggc aacattctct tctggaggcc 9660

gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggc cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcyccg ggtcgatgag acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatggc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgccccg gcaactcgtc 9960

gagggcaaag gaatagagta gatgccgacc gggggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agtcocggag 10080

ttgagacaaa tgggtgttcg gatctcgata agatcgttc atttgtccaa gcagcaaaga 10140

gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttccgggtta 10200

ccctctccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtacgcy 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttccgggagc 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaato cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500

tcttccttga actctcaagc ctacaggaca cacattcatc gtaggataaa acctcgaaat 10560

canttctac taagatggta tacaatagta accatgcagc gttgcctagt gaatgctccg 10620

taacacccaa taocgcgggc gaaacttttt tacaactctc ctatgagtcg tttaccagaa 10680

atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740

gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagett ctacogcttg 10800

gaccagtcca gttccaggac cagtcgctcc acctcctcgg cgcacactgg aatgtgctgc 10860

agctcctgtg gacccaagaa cataccccag ggcgcgcac cgtacttgcc gctgtggtgt 10920

agctggtggg ccaactgtcag gcgcttcabg tagggcaggc cagcgatggg cccggtggga 10980

aagcgctgtg gcaccaggcc atcgtgtaca aacatatatg ccattgcgta tagcgtgatg 11040

cccagcccg ctcacaayca ygcgcgcgc aggaegtggg gcagccagaa gcaaaaggta 11100

cacaggagca tggcgggag tccattgatg attgcaaca agtcgttgcc ttcaaaaggt 11160

ccagtgcgag gtgtgtggtg gctcttgctc agcagccagc ccaggaggca ctcatgccag 11220

atggctttgt gtgcatacgc ggcatacatc tccatgcga gcgcgcaccc aaccaccaag 11280

aggagagtgc cagccacttc accccatggc actgcgcgc ccacggatcat gtgcattgca 11340

aatctcaggt agtgggcgaa gatggcaatg cctgacacgc caattgatgc tgcaatggcg 11400

gcagcctggt atgacagctg ctcccgcttg cgcggggcac gacgctctgc gatagccggg 11460

tcaagctgct ggagtgtcac atcggcgctg tgctcatcgc cgcgcgcgc agcctgcacg 11520

gttccacgcg cctcctctgt ctgtggtgct gccactcgca gccgaactaa cgagcaccgc 11580

tgagcatgca ggcagacttt gggcgcgctg atgtcgggg ctagtccaac cgcgggggcc 11640

ttgacgtgta ttgactgcag cttcgacagc atagagataa aataaaaaga gaagaaaaga 11700

aagtttgtac aatttttttt tgtttatata acataacgc tatgtcaaca tttagaataa 11760

gggggaaaaa atcttcacat atattcgaat gcacaagatt attttttgt tcgcttttt 11820

tggtcgggto atcgagattt agagtgtaat caaagatact gtcactcga gagcgttgca 11880

caggctgctg tttgcaaat tggatgtttg ccgaattagt aaaatacgca agcatttctt 11940

acctttccgc tcccttttcc taattctccc aaagactaaa tgaggaaaga taaaggacaa 12000

agaaaatgta aagacaaaga aattgaaac gatataaact tgcagcactg aagaccanaag 12060

caaattggta actattcttg tgtacaaaca tgtataaaaa aaaacttttt ttgtctcctg 12120

gaggacaaaa tttcaaactc cttgaagaag attgcttgta tatctatcat atgcatatat 12180

catatcgatg gaaaaagaaa gtcaggcatg tatttataaa aagaagaatg tgccatgctt 12240

ccgaatttct tttcacttct ttttcottat ctattttaat ctcaagcttg gcgtaatcat 12300

ggcatagct gtttctctg tgaaattyll atccgctcac aattccacac aacatacgag 12360

ccgaagcat aaagtgtaaa gctgggggtg cctaatgagt gagctaactc acattaattg 12420

cggtgcgctc actgcccgct ttccagtcgg gaaacctgtc gtgcagctg cattaatgaa 12480

tcggccaacg cgcggggaga ggcggtttgc gtattgggcc aaagacaaaa gggcgacatt 12540

caacgattg agggaggga ggttaattatt gacggaaatt attcattaaa ggtgaattat 12600

caccgtcacc gacttgagcc atttggaat tagagccagc aaaatcacca gtacaccat 12660

taccattagc aaggccggaa acgtcaccaa tgaaccatc gatagcagca ccgtaatcag 12720

tagcgacaga atcaagtttg cttttagcgt cagactgtag cgcgttttca tcggcatttt 12780

cggcatagc ccccttatta gcgtttgcca tottttcata atcaaaatca ccggaaccag 12840

agccaccacc ggaaccgctt cctcagagc cgccaccctc agaaccgcca cctcagagc 12900

caccaccctc agagccgcca ccagaaccac caccagagcc gccgccagca ttgacaggag 12960

gcccgatcta gtaacataga tgacacggcg cgcgataatt tatcttagth tcgcgcctat 13020

attttgttt ctatcgctga ttaaatgtat aattgcggga ctctaactat aaaaaccat 13080

ctcataata acgtcatgca ttacatgtta attattacat gcttaacgta attcaacaga 13140

aattatatga taatcatcgc aagaccggca acaggattca atcttaagaa actttattgc 13200
caaatgtttg aaogatccgg gatcatccgg gtctgtggcg ggaactccac gaaaatatcc 13260
gaacgcagca agatatccgg gtgcacatcg gtcttgcccg ggcagtgccc gccagccgcg 13320
ttgatgtgga cgcggggccc gatcatattg tcgctcagga tcgtggcggt gtgcttgctg 13380
gcogttgctg tcgtaatgat atcgccacct tcgaccgccc gtccgcaga gatcccggtg 13440
gcgaagaact ccagcatgag atccccgcgc tggaggatca tccagccggc gtccccgaaa 13500
acgattccga agcccaacct ttcatagaag gcggcggtgg aatcgaaatc tcgtgatggc 13560
aggttgggcg tcgcttgctc ggtcatttcg aaccacagag tcccgctcag aagaactcgt 13620
caagaaggcg atagaaggcg atgcgctgcg aatcgggagc ggcgataccg taaagcacga 13680
ggaagcggtc agccattcgg ccgccaagct cttcagcaat atcacgggta gccaacgcta 13740
tgtcctgata gcggctccgc acaccacgac gccacagtc gatgaatcca gaaaagcgcg 13800
cattttccac catgatattc ggcgaagcag catcgccatg ggtcacgacg agatcatcgc 13860
cgtcgggcgt gcgcgccttg agcctggcga acagttcggc tggcgcgagc ccctgatgct 13920
cttcgtccag atcatcctga tcgacaagac cggcttccat ccgagtcagc gctcgctcga 13980
tcgatgtttt cgcttgggtg tcgaatgggc aggtagccgg atcaagcgta tgcagccgcc 14040
gcattgcato agccatgatg gatactttct cggcaggagc aaggtgagat gacaggagat 14100
cctgccccgg cacttcgccg aatagcagcc agtcctctcc cgcttcagtg acaacgtcga 14160
gcacagctcg gcaaggaaag cccgtcgtgg ccagccacga tagccgcgct gcctcgctct 14220
gcagttcatt cggggcaagc gacaggtcgg tcttgacaaa aagaaccggg cgcctcgtcg 14280
ctgacagccg gaacacggcg gcacagagc agccgattgt ctgttgctgc cagtcacagc 14340
cgaatagcct ctccaccaca gcggccggag aacctgcgtg caatccatct tgttcaatca 14400

tgcgaacga tccagatccg gtgcagatta tttggattga gagtgaatat gagactctaa 14460
ttggataccg aggggaattt atggnaacgtc agtggagcat ttttgacaag aaatatattgc 14520
tagctgatag tgaccttagg cgacttttga acgcgcaata atggttttctg acgtatgtgc 14580
ttagctcatt aaactccaga aaccgcggc tgagtggctc cttcaacgtt gcggttctgt 14640
cagttccaaa cgtaaaacgg cttgtccgc gtcacggcg ggggtcataa cgtgactccc 14700
ttaattctcc gctcatgac agattgtcgt ttccgcctt cagtttaaac tatcagtgtt 14760
tgacaggata tattggcggg taaacctaa agaaaagagc gtttattaga ataatcggat 14820
atttaaaagg gcgtgaaaag gtttatccgt tcgtccattt gtatgtgcat gccaacaca 14880
gggttcccca gatctggcg cgccagcga gacgagcaag attggccgcc gcccgaaacg 14940
atccgacagc gcgccagca caggtgcga ggcaattgc accaacgcat acagcgccag 15000
cagaatgcca tagtggggg tgacgtggtt agagtgaacc agatcgcgca ggaggcccg 15060
cagcacccgc ataatacaqc ccatgccgac agcgtcgagc gcgacagtgc tcagaattac 15120
gatcaggggt atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg 15180
aacgcgcgga ttctttatca ctgataagtt ggtggacata ttatgtttat cagtataaa 15240
gtgtcaagca tgacaaagt gcagccgaat acagtgatcc gtgcgcgccg ggacctgtt 15300
aacgaggtcg gcgtagacgg tctgacgaca cgcaactgg cggaacgggt gggggttcag 15360
cagccggcgc tttactggca cttcaggaa aagcgggcgc tgctcgagc actggccgaa 15420
gccatgctgg cggagaatca tacgcattcg gtgcgagag ccgacgacga ctggcgctca 15480
ttttgatcg ggaatgcccg cagcttcagg caggngctgc tcgcctaccg cgatcgccgc 15540
cgcacatcat ccggcacgac accgggcgca ccgcagatgg aaacggccga cgcgcagctt 15600
cgcttctct gcgaggcggg tttttcgcc ggggacgcg tcaatgcct gatgacaatc 15660

agctacttca ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgcggcgag 15720
cgcgggcgca ccgttgaaca ggctccgctc tcgcgcgtgt tgcggggcgc gatagagccc 15780
ttcgacgaag ccggtcocga cgcagcgttc gagcaggggac tcgcggtgat tgtcagtgga 15840
ttggcgaaaa ggaggtcgtt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga 15900
tcaggaccgc tgcgggagcg caaccactc actacagcag agccatgtag acaacatccc 15960
ctcccccttt ccaccgcgtc agacgcocgt agcagccgcg tacgggcttt ttcatgcocct 16020
gcocctagcgt ccaagcctca cggccgcgct cggcctctct ggccggccttc tggcgctctt 16080
ccgcttcttc gctcaactgac tcgctgcgct cggctggttc gctgcggcga gcgggtacag 16140
ctcactcaaa ggcggttaata cggttatcca cagaatcagg ggataacgca ggaagaaca 16200
tgtgagcaaa aggccagcaa aagccagga accgtaaaaa ggccgcgttg ctggcggttt 16260
tccataggct ccgcocccct gacgagcctc acaaaaatcg acgctcaagt cagagggtggc 16320
gaaaccgcag aggactataa agataccagg cgtttccccc tggaaagctcc ctctgtgcgt 16380
ctcctgttcc gacctgcgc cttaccggat acctgtccgc cttctctccc tcgggaagcg 16440
tggcgctttt ccgctgcata accctgcttc ggggtcatta tagcgatttt ttggtatct 16500
ccatcccttt tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg 16560
tgtatccaac ggcgtcagcc gggcaggata ggtgaagtag gcccaaccgc gagcgggtgt 16620
tcctttctca ctgtccctta ttgcacctg gcggtgctca accggaatcc tgctctcgca 16680
ggctggccgg ctaccgcggc cgtaacagat gaggggcaagc ggtgggtga tgaaaccaag 16740
ccaaccagga agggcagccc acctatcaag gtgtactgcc ttccagagca acgaagagcg 16800
attgaggaaa agggcgggcg gccggcctag agcctgtcgg cctacctgct ggcggtcggc 16860
cagggtcaca aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcctc 16920

aatggcgacc tgggcccgcct gggcggcctg ctgaaactct ggctoaccga cgaccgcgcg 16980
acggcgcggt tcggtgatgc cagatcctc gccctgctgg cgaagatcga agagaagcag 17040
gacgagcttg gcaaggtcat gatgggcgtg gtccgccga gggcagagcc atgaactttt 17100
tagccgctaa aacggccggg ggggtgcgct gattgccaag cagtcacca tgcgctccat 17160
caagaagagc gacttcgogg agctggtgaa gtacatcac cagcagcaag gcaagaccga 17220
gcgcctttgc gacgctca 17238

<210> 39
<211> 17238
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (10264)..(10264)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10472)..(10472)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10563)..(10563)
<223> n is a, c, g, or t

<400> 39
ccgggctggt tgcctcgcc gctgggctgy cggccgteta tggccctgca aacgcgccag 60
aaacgcgctc gaagcgtgt gcgagacac cgggccgccc gcgttgtgga taactcgccg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180
ccggcgcgcc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240
cagcctcgca aatcgcgcaa aacgcctgat ttacgcgag ttccccacag atgatgtgga 300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360
gagggggcgg atccttgaca ctgagggggc agagtgtcta gagatgaggg gcgcacatat 420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgaagggt ttccgcccgt 480
ttttcggcca ccgctaacct gtcttttaac ctgctlllca accaatatbt ataaacctg 540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600
cttctcgaa cctcccgccc cgctaaccgg ggcctcccat ccccccaggg gctgcgcccc 660
tcggcgcgca acggcctcac ccaaaaaatg gcagcgctgg cagtcttgc cattgcgggg 720
atcggggcag taacgggatg ggcgatcagc ccgagcgaga cgcccggaag cattgacgtg 780
ccgcaggtgc tggcatcgac attcagcgac caggtgcggg gcagtgaggg cggcgggcctg 840
ggtggcgccc tgcccttcac ttccggcgctc ggggcattca cggacttcat gggggggccc 900
gcaattttta ccttgggcat tcttggcata gtggtcggg gtgcctgtct cgtgttcggg 960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattatacgg aggtatgaaa 1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080
acgaagagga tgaagaggat gaggggcagg attgccttga atatattgac aatactgata 1140
agataatata tcttttatat agaagatctc gcggtatgta aggatttcag ggggcaaggc 1200
ataggcagcg cgcttatcaa tatacttata gaatgggcaa agcataaaaa ctgcatgga 1260
ctaagtcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac ttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgcag gtgtgcctc 1440

agattoaggt tatcgcgctc aattcgctgc gtatategct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccattccgta tccatatcac caggtcaaag 1560

ggtgacagca ggctcataag acgcccacgc gtccgcatag tgcgttccac gaatacgtgc 1620

gcaacaacgc tottccggag actgtcatac gcgtaaaaca gccagcgctg ggcgcgattta 1680

gccccgacat agccccactg ttcggtccatt tcgcgcgaga cgatgacgtc actgcocggc 1740

tgtatgcgcg aggttacoga ctgcggccty agltttttaa gtgacgtaaa atcggtgttga 1800

ggccaacgcc cataatcgcg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtacgc ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgcg ttacgcacca 1980

ccocgtcagt agctgaacag gagggacagc tcatagacac aqaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100

aaatcggtgc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggtt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220

aattagcttc ttggggatc ttttaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaantgaga atatcacgcg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaaggata taagctgggt ggagaaaatg aaacdtata 2400

tttaaaaatg acggaccgcg ggtataaagg gaaccactat gatgtggaaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccgtcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc aaggctctt 2640
tcactccatc gacatatacg attgtcccta taogaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgaatggat tgcgaaaact gggaagaaga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
ggaacttgtc ttttccacg ccgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc ttattgatc ttgggagaag cggcagggcg gacaagtggc atgacattgc 2940
cttctcgctc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgaccgact 3120
tcttcgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggtcgtcgtt attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cggcttaagg gaccgaattc attgctgata aggtggatta tctggacacc aagccacca 3300
gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360
gagggtgaat gaatcggagc ttgacogga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc cggagtgacc gaaaccatcg caagccgcac cgtcatgcgt ggcggccgcg 3480
aaaccttcca gtccgtgggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540
gcgtgcaact ggtcccccct gccctgcccg cggcatgggc cggcgtggag cgttcgcgtc 3600
gtctcgaaac ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaaacta 3660
tgacgacca gaaagcaaaa accgcggcg aggaactggc aaaaacaggtc agcaggacca 3720
agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaatgcag ctttccttgt 3780
tcgatattgc gccgtggcg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac caccgcgaac aagaaaatcc cgcgcgagcg gctgcaaaac aaggtcatatt 3900
 tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
 acgaactggg gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggccagccga 4020
 tcacottcac gttctacgag ctttgccagg aactgggctg gtgatcaat gcccggtatt 4080
 acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac gccgatgggc ttacgtcccg 4140
 accgcgttgg gcacctggaa tcggtgtcgc tgcgtcaccg ctccgcgtc ctggaccgtg 4200
 gcaagaaaac gtcccgttgc caggtoctga tcgacgagga aatcgctgtg ctgtttgctg 4260
 gccaccacta cagcaaatcc atatgggaga agtaccgcaa gctgtgcgcc accgcccagc 4320
 ggatgttoga ctatttcagc tcgcaocggg agccgtaccc gctcaagctg gaaaccttc 4380
 gcctcatgtg cggatcggat tcaccccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
 cctgcgaaga gttgcgagcg agcggcctgg tggaaacacg ctgggtcaat gatgacctgg 4500
 tgcattgcna acgctagggc ctttgtgggt cagttccggc tgggggttca gcagccagcg 4560
 ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gtcagtatc 4620
 gctcgggacg caccggcgcg tctacgaact gccgataaac agaggattaa aattgacaat 4680
 tgtgattaag gtcagattc gacggccttg agcgccgac gtgcaggatt tcgcgagat 4740
 ccgattgtcg gccctgaaga aagctccaga gatgttcggg tcggtttacg agcacgagga 4800
 gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gcctggcat tcggcgccca 4860
 catcgacggd gagatcattg ggctgtcggt cttcaaacag gaggacggcc ccaaggacgc 4920
 tcacaaggcg catctgtccg gcgttttngt ggagcccga caccagggcc gaggggtcgc 4980
 cggtatgtgt ctgcgggcgt tgccgggggg ttattgtctc gtgatgatcg tccgacagat 5040
 tccaacggga atctggtgga tcgcacatct catctcggc gcacttaata ttctgctatt 5100

ctggagcttg ttgtttattt cggctaccg cctgcgggc ggggtgcgg cgacggtagg 5160
cgctgtgcag ccgtgatgg tcgtgttcac ctctgcgct ctgctaggta gcccgatacg 5220
attgatggcg gtcctggggg ctatttgccg aactgcgggc gtggcgctgt tgggtttgac 5280
accaaacgca gcgctagatc ctgtggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgctcgga acgctgctga ccgcgaagt gcaacctccc gtgcctctgc tcacctttac 5400
cgctcggcaa ctggcgcccg gaggacttct gctcgttcca gtacgtttag tgtttgatcc 5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg cgttggtcgc gcctgatcgg 5520
agcgggttta acctactcc tttggttccg ggggatctcg cgactcgaac ctacagtgtg 5580
ttccttactg ggcttttcca gcccagatc tggggtcgat cagccgggga tgcacaggc 5640
cgacagtcgg aacttcgggt cccgcacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcctc aacgttccact tctaaagaaa tagcgcact cagcttctcc agcggcttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt caccgttaag 5820
cgagaaatga ataagaagcg tgataattcg gatctctcgc agggagatga tatttgatca 5880
caggcagcaa cgtctctgta tggttacaat caacatgcta cctccgcga gatcactgt 5940
gtttcaaacg cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaa 6000
tctgcgcctc tacaacggct ctcccgctga cgcgctcccg gactgatggg ctgcctgtat 6060
cagatgggtg ttttgtccg agctgcgggt cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtggttt tcttttcacc agtgagacgg gcaacagctg attgccttc 6240
accgcctggc cctgagagag ttgcagcaag cgggtccacg tggtttgccc cagcaggcga 6300
aaatcctggt tgatggtggt tccgaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
 actocaaact caaagggcga aaaacccgtct atcagggcga tggccacta cgtgaacct 6480
 caccctaaac aagttttttt gggctcgaggt gccgtaaagc actaaatcgg aacctctaaag 6540
 ggagccccc atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600
 agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgattcggg 6660
 cgggcctctt cgtctattac ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
 tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgaaggccag tgaattcgag 6780
 ctcggtaccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840
 cgaggagcct cgtctctgtc caactaccaa catggagtac gataagggcc agttccgcc 6900
 gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgaacttcaa 6960
 gtacaccaac gctctcttga tccagtcgat catccgtga aggcgcttc gaactcgggt 7020
 aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080
 aacagcttcc tcagccagg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140
 aagaactgga ggggtgtgt caaggaggag taagctctt attgaagtc gaggcggag 7200
 cgggtgcaag aggatattct tgcactctgt attatagata agatgatgag gaattggagg 7260
 tagcatagct tcatttggat ttgctttcca ggctgagact ctgcttgga gcatagagg 7320
 tcttttggtt ttcaatatcc tcaagtatct cgagtttgaa cttattccct gtgaacctt 7380
 tattaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt 7440
 tcgaaataca tccggtatgc gaaggcttgg ggcacctcgc ttggttgaat ttgaacgtg 7500
 gcactattga tcatcogata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560
 tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac ctctagtgga ctctagctacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactatata gaataagata ggtggagago ttataccgag ctoccaaate 7920

tgccagatc atggttgacc ggtgcctgga tcttctata gaatcatctt tattcgttga 7980

cctagctgat tctggagtga ccagaggggt catgacttga gccaaaaac cgcgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaacgg tgactctttc 8100

tggcatgcgg agagacggac ggcgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggacgg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactattt gcatcatcgg 8280

agaatatgga gcttcacga atcacggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgccg 8460

gcgtaaagtc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cggaacaca agctggcagt cgaccatcc ggtgctctgc actcgacctg ctgaggtecc 8640

tcagtcctcg gtaggcagct ttgcccgtc tgcgcgccg gtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtccaa catttgttgc catattttcc tgctctcccc accagctgct 8760

cttttttttt ctcttttttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttccctt aagtaagtac ttgtctacat ccatactcca tcttcccatc ccttatttcc 8880

tttgaacctt tcagttcgag ctttcccaact tcatcgcagc ttgactaaca gctaccocgc 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagtccgac agcgtctccg aacctgatga gctctcggag ggcgaagaat ctcgctcttt 9060

cagcttcgat gtagaggggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaazgat cgttatgttt atcggcactt tgcctcggcc gcgctccoga ttocggaagt 9180

gcttgacatt ggggaattca gcgagagcct gaacctattgc atctcccgcc gtgcacaggg 9240

tgctcagttg caagacctgc ctgaaccucca actgcccggc gttctgcagc cggctcggga 9300

ggccatggat gcgctcgcgt cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480

tctcgtatgg ctgatgcttt gggccgagga ctgccccgaa gtcgggcacc tctgcacgc 9540

ggatttcggc tccaacaatg tctgcagga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attccaata cgaggtcgc aacatcttet tctggaggcc 9660

gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgcgc cggctccggg cgtatatgct ccgcattggg cttgaccaac tctatcagag 9780

cttgggttgc ggcgaatttc atgatgcgc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gcggggactg tcggggctac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggo tgtgtagaag tactgcoga tagtggaacc cgaacccca gcaactcgtc 9960

gagggcaazg gaatagagta gatccgcgc ggggatcga tccacttaac gttactgaaa 10020

tcatacaaca gcttgacgaa tctggatata agatcgttgg tctcgatgic agctccggag 10080

ttgagacaaa tgggtttcag gatctogata agatacgttc atttgcctaa gcagcaaaaga 10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcggtt ttccgggttta 10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagttaacgc 10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttccgggagac 10320
gagatcaacg agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgtctttggc 10380
tccacggcgc tatatatttg tctctaattg tactttgaca tgcctctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaagata attgcatggt 10500
tcttcttga actctcaacg ctacaggaca cacattcatc gtagggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgcgggcc gaaacttttt tacaactctc ctatgagtcg tttaccoga 10680
atgcacagggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctgctgtact 10740
gtgttaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt agagataaaa 10800
taaaaagaga agaaaagaaa gtttgtacaa tttctttttg tttatataac atacacgcta 10860
tgtcaacatt tagaataagg gggaaaaaat cttccatcat attcgaatgc acaagattat 10920
ttctttgttc gctctttttg gtccgggtcat cgagatttag agtgtaatca aagatactgt 10980
catctcgaga gcgttgaca ggctgctgtt tgccaaattg gatgtttgca gaattagtaa 11040
aatacgcaag catttcttacc ctttccgctc ccttttcta attctcccaa agactaaatg 11100
aggaaagata aaggacaaag aaaatgtaaa gacaaagaaa ttgaaaacga tataaacttg 11160
cagcacgtaa gaccaaagca aattggtaac tatctttgtg tacaacatg tataaaaaaa 11220
aacttttttt tgctctcgga ggacaaaatt tcaaaactct tgaagaagat tgctgtgata 11280
tctatcatat gcataatca tatcgatgga aaaagaaagt caggcatgta tttataaaaa 11340

gaagaatgtg coatgettoc gaattttttt tcactttctt ttccttatct attttaatct 11400
catgetgtcg aagctgcagt caatcagcgt caaggcccg cgcgttgaac tagccgcga 11460
catcagcgg cccaaagtct gcctgcctgc tcagcgggtc tegttagtgc ggctgcgagt 11520
ggcagcacca cagacagagg aggcgctggg aaccgtgcag gctgcggcg cggcgcatga 11580
gcacagcgcc gatgtagcac tccagcagct tgacggggct atcgagagc gtcgtgccg 11640
gcgcaaacgg gagcagctgt cataccagcg tcccgccatt gcagcatcaa ttggcgtgtc 11700
aggcattgcc atcttcgcca cctacotgag atttgccatg cacatgaccg tggcgggcgc 11760
agtgcctatg ggtgaagtgg ctggcactct cctcttgggtg gttggtggcg cgcctggcat 11820
ggagatgtat gccgcctatg cacacaaagc catctggcat gactgcctc tgggctggct 11880
gtgcacaag agccaccaca cacctgcac tggacccttt gaagccaaag acttgtttgc 11940
aatcatcaat ggactgcgc coactgetct gtgtacctt ggcttctggc tgcccaact 12000
cctggggggc gctgtcttg gagcgggct gggcatcag ctatacgga tggcatatat 12060
gtttgtacac gatggcctgg tgcacaggcg ctttccacc gggcccatcg ctggcctgcc 12120
ctacatgaag gcctgcagc tggccacca gctacaccac agcggcaagt acggtggcgc 12180
gccctgggtg atgttcttgg gtccacagga gctgcagcac attccaggtg cggcgaggga 12240
ggtggagoga ctggtcctgg aactggactg gtccaagcgg tagaagcttg gcgtaatcat 12300
ggtcatagct gtttccctgt tgaattgtt atccgtcac aattccacac aacatacag 12360
cgggaagcat aaagtgtaaa gcctgggggtg cctaagtgt gagctaacto acattaattg 12420
cgttgcgctc actgcgcct ttcagtcgg gaaactgtc gtgccagctg cattaatgaa 12480
tcggccaacg cgcggggaga ggcggtttgc gtattgggcc aaagacaaa gggcgacatt 12540
caacogattg agggaggga ggtaaatatt gacggaatt attcattaaa ggtgaattat 12600

caccgtcacc gacttgagcc atttggaat tagagccagc aaatcacca gtgacccat 12660
taccattagc aaggccggaa acgtcaccaa tgaaccatc gatagcagca cgtaatcag 12720
tagcgacaga atcaagtttg cctttagcgt cagactgtag cgcgttttca tcggcatttt 12780
cggtcatagc ccccttatta gcgtttgcca tcttttata atcaaaatca cgggaaccag 12840
agccaccacc ggaaccgcct ccctcagagc cgcaccctc agaaccgcca cctcagagc 12900
caccaccctc agagccgcca ccagaaccac caccagagcc gccgcagca ttgacaggag 12960
gccgatcta gtaacataga tgacaccgcg cgcgataatt tatcctagtt tgcgcgctat 13020
atttgtttt ctatcgcta ttaaatgtat aattgcggga ctctaatcat aaaaacctat 13080
ctcataata acgtcatgca ttacatgtta attattacat gcttaacgta attcaacaga 13140
aattatatga taatcatgc aagaccggca acaggattca atcttaagaa actttattgc 13200
caaatgttg aacgatcgg gatcatcgg gtctgtggcg ggaactccac gaaaatatcc 13260
gaacgcagca agatatcgg gtgcacatcg gtcttacctg ggcagtcgac gccgcagccg 13320
ttgatgtgga cgcggggccc gatcatattg tcgctcagga tcgtggcggt gtgcttgcg 13380
gcggttgctg tcgtaatgat atcggcacct tcgaccgcct gtccgcaga gatcccgtyg 13440
gcgaagaact ccagcatgag atccccgcgc tggaggatca tccagccggc gtcccggaaa 13500
acgattccga agcccaacct ttcatagaag ggcgggtgg aatcgaaato tcgtgatggc 13560
aggttgggag tcgcttggtc ggtcatttcg aacccagag tccgcctcag aagaactcgt 13620
caagaaggcg atagaaggcg atgcgctgcg aatcgggagc ggcgataccg taaagcacga 13680
ggaagcggto agcccattcg ccgccaagct ctccagcaat atcacgggta gccaacgcta 13740
tgtctgata gcggtccgcc acaccagcc ggcacagtc gatgaatoca gaaagcggc 13800
cattttcac catgatattc ggcaagcagg catcgcatg ggtcacgagc agatcatgc 13860

cgtcgggcat gcgcgccttg agcctggcga acagttcggc tggcgcgagc cctcgatgct 13920
cttcgtccag atcatcctga tcgacaagac cggcttccat ccgagtagct gctcgtctga 13980
tgcgatgttt cgcttggtgg tcgaatgggc aggttagcgg atcaagcgta tgcagccgcc 14040
gcattgcctc agccatgatg gatactttct cggcaggagc aaggtagatg gacaggagat 14100
cctgcceccg cacttcgccc aatagcagcc agtcccttcc cgcttcagtg acaacgtcga 14160
gcacagctgc gcaaggaaac cccgtcgtgg ccagccacga tagccgcgct gcctcgtcct 14220
gcagttcatt caggggcaccg gcaayylcgg tcttgacaaa aagaacccgg cgcctcgtcg 14280
ctgacagccg gaacacggcg gcacagagc agccgattgt ctgttgtgccc cagtcacagc 14340
cgaatagcct ctccacccaa gcggccggag aaactgcgtg caatccatct tgttcaatca 14400
tgcgaaacga tccagatccg gtgcagatta tttggattga gagtgaatat gagactctaa 14460
ttggataccg aggggaattt atggaacgtc agtgagcat ttttgacaag aaatatattgc 14520
tagctgatag tgaccttagg cgacttttga acgcgcaata atggtttctg acgtatgtgc 14580
ttagctcatt aaactccaga aaccgcggc tgagtggctc ctcaaacgtt gcggttctgt 14640
cagttccaaa cgtaaaacgg ctgttccccc gtcacggcg ggggtcataa cgtgactccc 14700
ttaattctcc gctcatgac agattgtcgt tccccgctt cagtttaaac tatcagtggt 14760
tgacaggata tattggcggg taaacctaa agaaaagagc gtttattaga ataactggat 14820
atttaaaagg cgtgaaaag gtttatccgt tcgtccattt gtatgtgcac gccaacca 14880
gggttcccca gatctggcg cggccagcga gacgagcaag attggccgcc gcccgaaacg 14940
atccgacagc gcgcccagca cagggtgcga ggcaaattgc accaacgcac acagcgccag 15000
cagaatgcca tagtggggcg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg 15060
cagcacggcg ataactcagg cgatgcgcac agcgtcgagc gcgacagtc tcagaattac 15120

gatcaggggt atgttgggtt tcaagtctgg cctccggacc agcctccgct ggtccgattg 15180
aacgcgcgga ttctttatca ctgataagtt ggtggacata ttatgtttat cagtataaaa 15240
gtgtcaagca tgacaaaagt gcagccgaat acagtgatec gtgccgccct ggacctgttg 15300
aacgaggtcg gcgtagacgg tctgacgaca cgcaactgg cggaacggtt gggggttcag 15360
cagccggcgc tttactggca cttcaggaac aagcgggcgc tgcctgacgc actggccgaa 15420
ggcatgctgg cggagaatca tacgattcgt gtgccgagag ccgacgacga ctggcgctca 15480
tllctgacg ggaatgcccc cagcttcagg caggcgctgc tcgcctaacg cgatggcgcg 15540
cgcatccatg ccggcacgcg accggggcga ccgcagatgg aaacggccga ccgcagcctt 15600
cgcttctct gcgaggcggg tttttcgcc ggggacgcgc tcaatgcgct gatgacaatc 15660
agctacttca ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgcggcgag 15720
cgcggcggca ccgttgaaca ggctccgctc tcgccgctgt tcggggccgc gatagacgcc 15780
ttcgacgaag ccggtccgga cgcagcgttc gagcagggac tcgcggtgat tgcgatgga 15840
ttggcgaaaa ggagctcgt tgtcaggaaac gttgaaggac cgagaaaggg tgacgattga 15900
tcaggaccgc tgccggagcg caaccactc actacacgag agccatgag acaacatccc 15960
ctccccctt ccacgcgctc agacgccgct agcagccgcg taagggtctt ttcatgcct 16020
gccctagcgt ccaagcctca cggccgcgct cggcctctct ggccgccttc tggcgtctt 16080
ccgttctct gctcactgac tcgctgcgct cgggtgttcg gctgcggcga gcggtatcag 16140
ctcactcaaa ggcgtaata cggttatcca cagaatcagg ggataacgca ggaagaaca 16200
tgtgagcaaa aggcagcaa aaggccagga accgtaaaaa ggccgcgttg ctgggtttt 16260
tccataggct ccgccccct gacgagcctc acaaaaatcg accgtcaagt cagaggtggc 16320
gaaacccgac aggactataa agataccagg cgtttcccc tggaagctcc ctgctgcgct 16380

ctcctgttcc gaccctgcgg cttaccggat aactgtccgc ctttctccct tcgggaagcg 16440
tggcgctttt ccgtgcata aacctgcttc ggggtcatta tagcgatttt ttccgtatat 16500
ccatcctttt tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg 16560
tgtatccaac ggcgtcagcc gggcaggata ggtgaagtag gcccacccgc gagcgggtgt 16620
tcctttctta ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga 16680
ggctggccgg ctaccgcgg cgtaacagat gagggaagc ggatggctga tgaaccaag 16740
ccaaccagga agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg 16800
attgaggaaa agggggcgcc ggccggcatg agcctgtcgg cctacctgct ggccgtccgc 16860
cagggttaca aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc 16920
aatggcgacc tgggccgctt gggcgccctg ctgaaactct ggctcaccga cgaccgcgc 16980
acggcgcggt tcgggtgatc cagatccctc gcctgctgg cgaagatcga agagaagcag 17040
gacgagcttg gcaaggtcat gatggcgctg gtccgccga gggcagagcc atgacttttt 17100
tagccgctaa aacggccggg gggcgcgctt gattgccaag caggtcccca tgcgctccat 17160
caagaagagc gacttcggg agctgggtgaa gtacatcacc gacgagcaag gcaagacga 17220
gcgcctttgc gacgtca 17238

<210> 40

<211> 18449

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 40

gatcttttcga cactgaaata cgtcgagcct gctocgcttg gaagcggcga ggagcctcgt 60

cctgtcacia ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120

cagttcatgg gcgttggcat gatggccgtc atgcatttgt acttcaagta caccaacgct 180

cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggtttaag atccacgtct 240

tcgggaagcc agcgactggt gacctccagc gtocctttaa ggcgtccaac agctttctca 300

gccaggggcca gcccaagacc gacaaggcct ccctccagaa cgcgcgagaag aactggaggg 360

gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg 420

atatctcttg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca 480

tttgattttg ctttccaggc tgagactcta gcttggagca tagagggtcc tttgctttc 540

aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga 600

gcatttggat gaacatgaat ctgaggactg caatogccat gaggttttcc aaatacatcc 660

ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca 720

tccgatagct ctgcaagggg cgttgcacia tgcaagtcga acgttgctag cagttccagg 780

tggaaatgta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840
 caccacaaaa gtcagacggc gtaacacaaa gtcacacaa acaagctgta aggatttcgg 900
 caoggotacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt 960
 tgtgtttgat cgagacctaa tacagccctt acaacgacca tcaaagtcgt atagctacca 1020
 gtg^{aggaagt} gactcaaat cgaacttcagc aacatctcct ggataaactt taagcctaaa 1080
 ctatacagaa taagataggt ggagagctta taccgagctc ccaaatctgt ccagatcatg 1140
 gttgaccggt goctggatct tcctatagaa tcactcttat tcgttgacct agctgattct 1200
 ggagtgacc caggggtcat gacttgagcc taaaatccgc cgcctccacc atttgtagaa 1260
 aatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttctgg catgaggaga 1320
 gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380
 ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440
 aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcactcgaga atatggagct 1500
 tcactgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560
 cgaaatagca tgcattaac ctaggtagag aagccaatt gcttcgcatc tggtaaaga 1620
 ttacagagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680
 aattggocaa tccggcatct gtagggcgct caaatatcgt gctctcctg ctttgccegg 1740
 tgtatgaac oggaaaggcc gctcaggagc tggccagcgg cgcagacggg gaacacaagc 1800
 tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtccttgta 1860
 ggcagctttt ccccgctctg ccgccegggtg tctggcggg gttgacaagg tctgtgcgtc 1920
 agtccaaatc ttgttgcatc attttctctg tctcccccac agctgctctt tctttttctc 1980
 tttcttttcc catcttcagt atattcatct tcccatccaa gaacctttat tccccctaag 2040

taagtacttt gctacatcca tactccatcc ttcccatccc ttattccctt gaacctttca 2100
 gttcgagctt tcccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160
 accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220
 gtctcogacc tgatgcagct ctgggagggc gaagaatctc gtgctttcag cttcgatgta 2280
 ggaggggctg gatattgtct gcgggtaaat agctgcgcgc aggttttcta caaagatcgt 2340
 tatgtttatc ggcactttgc atcgcccgcg ctcccgattc cggaggtgct tgacattggg 2400
 gaattcagcg agagcctgac ctattgcato tcccgccgtg cacagggtgt caogttgcaa 2460
 gacctgctg aaacgaact gcccgctgtt ctgcagccgc tcgcggaggg catggatgcy 2520
 atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcgacc gcaaggaatc 2580
 ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640
 tggcaaaactg tgatggagca cacogtcagt gogtcogtcg cgcaggctct cgtagagctg 2700
 atgctttggg ccgaggactg ccccgaaatc cggcacctcg tgcacgcgga ttccggctcc 2760
 aacaatgtcc tgacggacaa tggccgcata acagcgttca ttgactggag cgaggcgatg 2820
 ttccggggatt cccaatacga ggtcgccaac atcttcttct gaggccgtg gttgcttgt 2880
 atggagcagc agacgcgcta cttcgagcgg aggcacccgg agcttgacgg atcgccgcgg 2940
 ctccggcgct atatgtccg cattggctct gaccaactct atcagagctt ggttgacggc 3000
 aatttcgatg atgcagcttg ggccgagggc cgtatgcagc caatcgtccg atccggagcc 3060
 gggactgtcg ggcgtacaca aatcgccgcg agaagcgcgg ccgtctggac cgtggctgt 3120
 gtagaagtac tcgccgatag tggaaaccga cgcgccagca ctcttcogag ggcaaggaa 3180
 tagagtagat gccagccggc ggatcgatcc acttaacggt actgaaatca tcaaacagct 3240
 tgacgaatct ggatataaga tcgttggtgt cgtatgcagc tccggagttg agacaaatgg 3300

tgttcaggat ctgcataaga tacgttcatt tgtccaagca gcaaagagtg ccttctagtg 3360
 atttaaatgc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420
 cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480
 gcgaagagaa gaatagetta gcagagctat ttctattttc gggagacgag atcaagcaga 3540
 tcaacggctg tcaagagacc tacgagactg aggaatccgc tcttggtctc acgcgactat 3600
 atatttgtct ctaattgtac ttgacatgc tctctctctt tactctgata gcttgactat 3660
 gaaaatttcc tccaccgccc ctgggttccc aaagataatt gcatgtttct tctctgaact 3720
 ctoaagccta caggacacac attcatcgta ggtataaacc tggaaatcan ttctactaa 3780
 gatggtatac aatagtaacc atgcatggtt gcctagttaa tgctccgtaa caccaatac 3840
 gcggccgcaa acttttttac aactctccta tgagtcgttt acccagaatg cacaggatca 3900
 cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960
 ctccacatct ccatcgcacc tgcaggcatg caaagcttga gattaaaata gataaggaaa 4020
 agaaagttaa aagaaattcg gaagcatggc acattcttct ttttataaat acatgcttga 4080
 ttttcttttt coactgatat gatatatgca tatgatagat atcaagcaa tcttcttcaa 4140
 ggagtttgaa attttgtcct ccaggagcaa aaaaagttt tttttatata atgtttgtac 4200
 acaagaatag ttaccaattt gctttggtct tacgtgctgc aagtttatat cgttttcaat 4260
 ttctttgtct ttacattttc ttgtccttt atctttctc atttagtctt tgggagaatt 4320
 aggaaaaggg agcggaaaagg taagaaatgc ttgcgtatct tactaattcg gcaaacatcc 4380
 aatttggcaa acagcagcct gtgcaacgct ctcgagatga cagtatcttt gattacacto 4440
 taaatctcga tgccccgacc aaaaagagcg aacaaagaaa taatcttgtg cattcgaata 4500
 Lgatgyaaga ttttttcccc cttattctaa atgttgacal agcgtgtatg ttatataaac 4560

aaaaagaaat tgtacaaact ttctttttctt ctcttttttat ttatatctcta tgctgtcgaa 4620
 gctgcagtc aacagcgtca agcccgcccg cgttgaacta gcccgcgaca tcacgcggcc 4680
 caaagtctgc ctgcagtctc agcgggtgctc gttagttcgg ctgcgagtg cagcaccaca 4740
 gacagaggag gcgctgggaa ccgtgcaggc tgcgggcgag ggcgatgagc acagcccgca 4800
 tgtagcactc cagcagcttg accgggctat cgcagagcgt cgtgcccgcc gcaaacggga 4860
 gcagctgtca taccaggctg ccgccattgc agcatcaatt ggcgtgtcag gcattgccat 4920
 cttcgccacc taactgagat ttgcctatga catgaccgtg ggcgcgcagc tgcctatggg 4980
 tgaagtggct ggcactctcc tcttggtggt tgggtggcgc ctggcatgg agatgtatgc 5040
 ccgctatgca cacaagcca tctggcatga gtgcctctg ggcgtggctgc tgcacaagag 5100
 ccaccacaca cctgcgactg gaccctttga agccaacgac ttgtttgcaa tcataatgg 5160
 actgcccgc atgtctctgt gtacctttgg ctctggtctg cccaacgtcc tgggggcggc 5220
 ctgctttgga gcggggctgg goatacgcct atacggcatg gcataatatgt ttgtacacga 5280
 tggcctggtg cacaggcgct ttcccaccgg gcccatcgct ggcctgcct acatgaagcg 5340
 cctgacagtg gcccccagc taccccacag cggcaagtao ggtggcgccg cctggggtat 5400
 gttcttgggt ccacaggagc tgcagcacat tccagggtcg gcggaggagg tggagcgact 5460
 ggtcctggaa ctggactggt ccaagcggtg gattgtgact gatagcgaga ctctgggtcg 5520
 atgttatctg cctcaacaat ggcttagaaa agaagaaaca gaacaaatac agcaaggcaa 5580
 cgcgcgtagc ctaggtagc aaagactggt ggccttgtct ctgaagcttg taggaaggcc 5640
 agacgctatc atggtgagag ctaagaaggc cattgacaag ttgccggcaa actgtcaagg 5700
 cgggtgtacga gctgcttgcc aagtatatgc tgcaattgga tctgtactca agcagcagaa 5760
 gacaacatat cctacaagag ctcatctaaa aggaagcgaa cgtgccaaag ttgctctggt 5820

gagtgtatac aacctctatc aatctgaaga caagcctgtg gctctccgc aagctagaaa 5880

gattaagagt ttttttgtg alaglyaal tttgtttt tttatgtcty atagttcaal 5940

aaagagacaa cacatacaat ataaaatcat tgtcttttaa tgtaattta gttagagtga 6000

aagcctgcat tttttttgta cgcataaaca atgaattcac cccgcttcgt gtttttaaat 6060

aattatgtca aactagggaa aattctttt tttctcttcg tctttttttt ggcttgttgt 6120

ggagtacacag gcttgtcttc agattgatag aggttgtata cactcaacag agcaatcttg 6180

gcacgttcgc ttctctttag atgagctctt gtaggatatg ttgtcttcgt ctgcttgagt 6240

acagatccaa ttgcagcata tacttggcaa gcagctcgta caccgccttg acagtttgcc 6300

ggcaacttgt caatgccctt cttagctctc accatgatag cgtctgcctt tcttacaagc 6360

ttcagagaca agcccaacag tctttgatca ctaggctac gggcgttgcc ttgctgtatt 6420

tgttctgttt cttcttttct aagccattgt tgaggcagat aacatcgacc caacatcttc 6480

gagccatact acagcataaa aggatacgtt ttctttaaca gaaatttacc cttttgttat 6540

cagcacatac aaaaaaaag aaatttaaga tgagtaggac ttccattctc tcaaaaaatt 6600

tattcaatcc ataatgaat tatttttga caaaaaagaa agattatgcc tgattttctc 6660

tatttttttt ttttttacia ctocaccaat actttctagc ccagcttgcc gtaatcatgg 6720

tcatagtctg ttctgtgtg aaattgttat ccgctcacia ttccacacia catagcagcc 6780

ggaagcataa agtgtaaagc ctgggggtgc taatgagtga gctaactcac attaatgctg 6840

ttgcgctcac tgcgcgttt ccagtcggga aacctgtcgt gccagctgca ttaattgaatc 6900

ggcaaacgcg cggggagagg cgttttgcgt attgggcaaa agacaaaagg gcgacattca 6960

accgattgag ggaggggaag taaatattga cggaaattat tcattaaagg tgaattatca 7020

ccgtcaccca cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta 7080

ccattagcaa ggcgggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta 7140

gcgcagaaat caagttttgcc tttagcgta gactgtagcg cgttttcacg ggcattttcg 7200

gtcatagccc ccttatttagc gtttgccatc ttttcataat caaaatcacc ggaaccagag 7260

ccaccaccgg aaccgcctcc ctcagagccg ccaccctcag aaccgccacc ctcagagcca 7320

ccaccctcag agccgcaccc agaaccacca ccagagcccg cgcagcatt gacaggaggg 7380

ccgatctagt aacatagatg acaccgogcg cgataattta tcctagtgtt cgcgctatat 7440

tttgttttct atcgcgtatt aaatgtataa ttgggggact ctaatcataa aaaccatct 7500

cataaataac gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa 7560

ttatatgata atcatcgcaa gaccggcaac aggattcaat cttaagaaac tttattgcc 7620

aatgtttgaa cgcacgggga tcattccgggt ctgtggcggt aactccagca aaatatccga 7680

acgcagcaag atatcgcggt gcattctgggt ctgtccctggg cagtgcgcgc cgcgcgcgtt 7740

gagtgtggacg ccggggccga tcattattgtc gctcaggatc gtggcggtgt gcttgtcggc 7800

cgttgcgtgc gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccggtggc 7860

gaagaactcc agcatgagat cccgcgcgtg gaggatcacc cagccggcgt cccggaaaac 7920

gattccgaag cccaaccttt catagaaggc ggcggtggaa tcgaaatctc gtgatggcag 7980

gttggcggtc gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca 8040

agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg 8100

aaagggtcag cccattcgcc gccaaactct tcagcaatat cagggttagc caacgctatg 8160

tcctgatagc ggtccgcacc accagccgg ccacagtcga tgaatccaga aaagcgcca 8220

ttttccacca tgatattcgg caagcagcca tcgccatggg tcacgacgag atcatcgccg 8280

tcgggcatgc gcgccttgag cctggcgaac agttcggctg gcgcgagccc ctgatgctct 8340

tcgtccagat catcctgac gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg 8400

cgatgtttcg cttgggtggtc gaatgggcag gtaccgggat caagcgtatg cagccgccgc 8460

attgcatcag ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc 8520

tgccccgcga cttgcgccaa tagcagccag tcccttcccg cttcagtgac aacgtcgagc 8580

acagctgcgc aaggaaacgc cgtcgtggcc agccacgata gcgcgctgc ctcgtcctgc 8640

agttcattca gggcacccga caggtcgggtc ttgacaaaa gaaccggcg cccctgcgct 8700

gacagccgga acacggcggc atcagagcag ccgattgtct gttgtgccca gtcatacgcc 8760

aatagcctct ccaccacaagc ggcggagaaa cctgcgtgca atccatcttg ttcaatcatg 8820

cgaacgac cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt 8880

ggataccgag gggaatttat ggaacgtcag tggagcattt ttgacaaaga atatttgcta 8940

getgatagtg accttagcgc acttttgaac gcgcaataat ggtttctgac gtatgtgctt 9000

agtcattaa actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca 9060

gttccaaacg taaaaaggct tgtccgcgt catcggcggg ggtcataacg tgactccctt 9120

aattctccg tcattgatcag attgtcgttt ccgccttca gtttaaaacta tcagtgtttg 9180

acaggatata ttggcgggta aacctagag aaaagagcgt ttattagaat aatcggatat 9240

ttaaaagggc gtgaaaaggc ttatccgttc gtccatttgt atgtgcacgc caaccacagg 9300

gttccccaga tctggcgccg gccagcgaga cgagcaagat tggccgcccgc ccgaaacgat 9360

ccgacagcgc gccacgaca ggtgcgcagg naaattgcac caacgcatac agcgcacgca 9420

gaatgccata gtggcggttg acgtcgttcg agtgaaccag atcgcgcagg aggcgccgca 9480

gcacggcat aatcaggccg atgcgcagag cgtcagcgc gacagtgcct agaattacga 9540

tcaggggtat gttgggttcc acgtctggcc tccggaccag cctccgctgg tccgattgaa 9600

cgcgccgatt ctttatoact gataagttgg tggacatatt atgtttatca gtgataaagt 9660

gtcaagcatg acaaagttgc agccgaatac agtgcacgtg gcgcgcctgg acctgttgaa 9720

cgaggtcggc gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca 9780

gccggcgctt tactggcact tcaggaacaa gggggcgctg ctgcacgcac tggccgaagc 9840

catgctggcg gagaatcata cgcattcggg gccgagagcc gacgacgact ggcgtcatt 9900

tctgacggg aatgcccgca gcttcaggca ggcctgctc ccctaccgag atgcgcgcgcg 9960

catccatgcc ggcacgcgac cgggcgcacc gcagatggaa acggccgacg cgcagcttcg 10020

cttctctgc gaggcggtt tttcgccggg ggacgcgctc aatgcgctga tgacaatcag 10080

ctacttcact gttggggcgg tgcttgaggga gcaggcgccg gacagcgatg ccggcgagcg 10140

cggcggcacc gttgaacagg ctccgctctc gccgctgttg cgggcgcgga tagacgcctt 10200

cgacgaagcc ggtccggacg cagcgttcga gcagggactc gcggtgattg togatggatt 10260

ggcgaagagg aggtctggtg tcagggaacgt tgaaggaccg agaaagggtg acgattgac 10320

aggacgcgctg ccggagcgca acccaactac tacagcagag ccattgtagac aacatccct 10380

ccccctttcc acgcgctcag acgcccgtag cagcccgcta cgggcttttt catgccctgc 10440

cctagcgctc aagcctcagc gccgcgctcg gcctctcttg cggccttctg gcgctcttcc 10500

gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgccggcagc ggtatcagct 10560

cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 10620

tgagcaaaag gccagcaaaa ggccagggaac cgtaaaaagg ccgcgttgct ggcgtttttc 10680

cataggctcc gcccccctga cgagcatcac aaaaatcgac gctcaagtca gagggtggca 10740

aaccgcagc gactataaag ataccaggcg ttccccctg gaagctccct cgtgcgctct 10800

cctgttccga cctgcgcgt taccggatac ctgtccgcct ttctcccttc ggggaagcgtg 10860

gcgcctttcc gctgcataac cctgcttcgg ggtcattata gcgatttttt cggatatatcc 10920

atcctttttc gcacgatata caggattttg ccaaaggggt cgtgtagact ttccttggtg 10980

tatccaacgg cgtcagccgg gcaggatagg tgaagtaggc ccaccgcga gcgggtgttc 11040

cttcttcaact gtcccttatt cgcacctggc ggtgctcaac gggaaatcctg ctctgcgagg 11100

ctggccggct accgcgggc taacagatga gggcaagcgg atggctgatg aaaccaagcc 11160

aaccaggaag ggcagcccac ctatcaaggt gtactccctt ccagacgaac gaagagcgat 11220

tgaggaaaag gcggcggcgg ccggcatgag cctgtcggcc tacctgctgg ccgtcgcca 11280

gggctacaaa atcacggggc tctgggacta tgagcacgtc ccgagctgg ccgcacatca 11340

tgcgacctg ggcgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgac 11400

ggcgcggttc ggtgatgcca cgatcctgc cctgctggcg aagatcgaa agaacgagga 11460

cgagcttggc aaggtcatga tgggcgtggt ccgccgagg gcagagccat gactttttta 11520

gcgcgtaaaa cggccggggg gtgcgcgtga ttgccaagca cgtccccatg cgtccatca 11580

agaagagcga ctctcgggag ctggtgaagt acatcacga cgagcaaggg aagacggagc 11640

gcctttgoga cgctcaccgg gctggttgcc ctgcgcgtg ggctggcggc cgtctatggc 11700

cctgcaaacg cgcgagaac gccgtcgaag ccgtgtgcga gacaccgcgg ccgccggcgt 11760

tgtggatacc tcgoggaaaa cttggccctc actgacagat gagggcgga cgttgacact 11820

tgagggggcg actcaccgg cgcggcgttg acagatgagg ggcaggctcg atttcggcg 11880

gcgacgtgga gctggccagc ctgcgaaac ggcgaaaac cctgatttta cgcgagtttc 11940

ccacagatga tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc 12000

gcgactactg acagatgagg ggcgggatcc ttgacactg aggggcgag tgctgacaga 12060

tgagggggcgc acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc 12120

aagggttttc gcccggtttt cgccacgcg taacctgtct tttaacctgc ttttaacca 12180

atatttataa accttgtttt taaccagggc tgcgcctgt gcgcgtgacc gcgcagccg 12240

aaggggggtg cccccccttc tcgaaccctc ccggcccgct aacgcgggac tcccatcccc 12300

ccaggggctg cgcgcctcgg ccgcgaacgg cctcaccoca aaaatggcag cgtctggcagt 12360

ccttgccatt gccgggatcg gggcagtaac gggatgggag atcagcccg gcgcgacgcc 12420

ccgaacgatt aacgtccccc aggtgctggc atcgacattc agcgaccagg tgcggggcag 12480

tgagggcggc ggcctgggtg gcggcctgcc cttcacttcg gccctgggg cattcacgga 12540

cttcattggc ggcgcggcaa tttttacctt gggcattott ggcatagttg tcgcgggtgc 12600

cgtgctcgtg ttgcgggggtg cgataaacc ccggaacat ttgaggtgat aggtaagatt 12660

ataccgaggt atgaaaacga gaattggacc ttacagaat tactctatga agcgccatat 12720

ttaaaaagct accaagcga agaggatgaa gaggatgagg aggcagattg ccttgaatat 12780

attgacaata ctgataagat aatatatctt ttatatagaa gatctcgccg tatgtaagga 12840

tttcaggggg caaggcatag gcagcgcgct tatcaatata totatagaat gggaaaagca 12900

taaaaacttg catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt 12960

ctatcataat tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgcc 13020

gatgactttg tcatgcagct ccacgcgatt tgagaacgac agcgacttcc gtcccagccg 13080

tgccagggtg tgccctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct 13140

gattactgtc agctttccct tcaggcggga ttcatacagc ggccagccat ccgtcatdca 13200

tatcaccaag tcaagggtg acagcaggct cataagcgc cccagcgctg ccatagtgag 13260

ttcacogaat acgtgogcaa caaccgtctt ccggagactg tcatacgctg aaaaacagca 13320

gcgctggcgc gatttagccc cgacatagcc ccactgttcg tccatttccg cgcagacgat 13380

gacgtcaactg ccgcgctgta tgcgcgaggt taccgactgc ggcctgagtt ttttaagtga 13440

cgtaaaatcg tgttgaggcc aacgccata atgcgggctg ttgcccgga tccaacgcca 13500

ttcatggcca tatcaatgat tttctgggtg gtaccgggtt gagaagcggg gtaagtgaac 13560

tgcaagtggc atgttttacg gcagtgaag cagagatagc gctgatgtcc ggcggtgctt 13620

ttgcggttac gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa 13680

gccactggag cacctcaaaa acaccatcat acactaaatc agtaagttag cagcatcacc 13740

cataattgtg gtttcaaat cggctccgtc gatactatgt tatacgccaa ctttgaatac 13800

aactttgaaa aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg 13860

gagttcgtct tgttataatt agcttcttgg ggatatctta aatactgtag aaaagaggaa 13920

ggaaataata aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat 13980

accgctcgct aaaagatacg gaaggaaatgt ctctgctaa ggtatataag ctggtgggag 14040

aaaatgaaaa cctatattta aaaatgacgg acagcgggta taaagggacc acctatgatg 14100

tggaaaggga aaaggacatg atgctatggc tggagggaaa gctgcctgtt ccaaagggtc 14160

tgcactttga acggcatgat ggctggagca atctgctcat gagtggggcc gatggcgtcc 14220

tttgctcgga agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg 14280

agtgcatacg gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag 14340

acagccgctt agcogaattg gattacttac tgaataacga tctggccgat gtggattggc 14400

aaaactggga agaagacact ccattttaag atccgcgcga gctgtatgat tttttaaaga 14460

cggaaaaagg cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct 14520

ttgtgaaaga tggcaaaagta agtggcctta ttgatcttgg gagaagcggc agggcggaca 14580

agtggtatga cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt. 14640

atgtcgagct atttttttgac ttactgggga tcaagcclga ttgggagaaa ataaaaalatt 14700

atatttttact ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag 14760

caggagcgca ccgacttctt ccgcatcaag tgttttggct ctcaggccga gggccacggc 14820

aagtatttgg gcaagggggt gctggtattc gtgcagggca agattcggaa taccaagtac 14880

gagaaggacg gccacagcgt ctacgggacc gacttcattg ccgataaggt ggattatctg 14940

gacaccaagg caccaggcgg gtcaaatcag gaataagggc acattgcccc ggcgtgagtc 15000

ggggcaatcc cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa 15060

gaactgatcg acgogggggt ttccgcgcag gatgccgaaa ccatcgcaag ccgcaccgtc 15120

atgcgtgcgc cccgcgaac cttccagtc gtccgctcga tggtcacga agctacggcc 15180

aagatcgagc gcgacagcgt gcaactggct cccctgccc tgcccgccgc atcgcccgcc 15240

gtggagcgtt cgcgtcgtct cgaacaggag gcggcagggt tggcgaagtc gatgaccatc 15300

gacacgcgag gaactatgac gaccaagaag cgaaaaaccg ccggcgaggga cctggcaaaa 15360

caggtcagcg agggccaagca ggcgcggttg ctgaaacaca cgaagcagca gatcaaggaa 15420

atgcagcttt ccttgttoga tattgcgcgc tggccggaca cgatgcgagc gatgccaaac 15480

gacacggccc gctctgcctt gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg 15540

caaaacaagg tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgtcgag 15600

ctcggggcgc acgatgacga actggtgtgg cagcaggtgt tggagtagc gaagcgcaac 15660

cctatcggcg agccgatcac cttcacgttc tacgagcttt gccaggacct gggcgtgctg 15720

atcaatggcc ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg 15780

atgggcttca cgtccgaccg cgttgggcac ctggaatcgg tgtcgtcgtc gcaccgcttc 15840
cgcgtcctcg accgtggcaa gaaaacgtcc cgttgccagg tctcgtatcga cagggaatc 15900
gtcgtgctgt ttgctggcga ccaactacacg aaattcatal gggagaayla ccgcaagctg 15960
tcgccgacgg ccgcacggat gttcgtactat ttcagctcgc accgggagcc gtaccgcctc 16020
aagctggaaa ccttcgcct catgtcggga tcggattoca ccgcgtgaa gaagtggcgc 16080
gagcaggctg gcgaagcctg cgaagagttg cagggcagcg gcctggtgga acacgcctgg 16140
gtcaatgatg acctggtgca ttgcaaacgc tagggccttg tggggctcagt tcggcctggg 16200
ggttcagcag ccagcgttt actggcattt caggacaag cgggcactgc tcgacgcact 16260
tgcttcgtc agtatcgtc gggacgcacg gcgcgtctca cgaactcgcg ataaacagag 16320
gattaaaatt gacaattgtg attaaggctc agattcgacg gcttgagcgc gccgacgtgc 16380
aggatttcgg cgagatccga ttgtcggccc tgaagaagc tcagagatg ttogggtccg 16440
tttacgagca cgaggagaaa aagcccatgg aggcgttcgc tgaacggtg cgagatgccg 16500
tggcattcgg cgcctacatc gacggcgaga tcattgggct gtoggctctc aaacaggagg 16560
acggccccaa ggaactcac aaggcgcac tgctcggcgt ttctgtggag ccgcaacagc 16620
gaggccgagg ggtcgccggt atgctgctgc gggcggtgac ggggggttta ttgctcgtga 16680
tgatcgtcgc acagattcca acgggaatct ggtggatgag catcttcac ctcggcgcac 16740
ttaatatctc gctattctgg agcttggtgt ttatttcggt ctaccgcctg ccggggcggg 16800
tcgcggcgac ggtaggcgct gtgcagccgc tgatggctgt gttcatctct gccgcctgc 16860
taggtagccc gatacgattg atggcggtcc tgggggctat ttcggaact gcgggcgtgg 16920
cgctgttggt gttacacca aacgcagcgc tagatcctgt cggcgtcgca gcgggcctgg 16980
cggggcggt ttocatggcg ttcggaaccg tctgacccg caagtggcaa cctcccgctc 17040

ctctgtctcac ctttaccgcc tggcaactgg cggccggagg acttctgtctc gttccagtag 17100

ctttagtgtt tgatccgcc atcccgatgc ctacaggaa caatgttctc ggctggcgt 17160

ggctcgccct gatcggaggg ggtttaacct acttcctttg gttccggggg atctcgcgac 17220

tcgaacctac agttgtttcc ttactgggct ttctcagccc cagatctggg gtgcgtacgc 17280

cggggatgca tcagccgcac agtcggaact tcgggtcccc gacctgtacc attcggtag 17340

caatggatag gggagttgat atcgtcaacg ttcaattcta aagaaatagc gccactcagc 17400

ttctcagcg gctttatcca gcgatttct attatgtcgg cataattctc aagatcgaca 17460

gcctgtcacg gtaaacgcag aaatgaataa gaaggctgat aattcggatc tctgcgggg 17520

agatgatatt tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgtaccct 17580

ccgcgagatc atccgtgttt caaacccggc agcttagttg ccgttcttcc gaatagcatc 17640

ggtaacatga gcaagctcg ccgccttaca acggctctcc cgctgacgcc gtcccgact 17700

gatgggctgc ctgtatcgag tggtgatttt gtgcgagct gccggtcggg gagctgttg 17760

ctggctggtg gcaggatata ttgttggtga aacaaattga cgttaggaca acttaataac 17820

acattcggga cgtttttaat gtactggggt ggtttttctt ttcaccagtg agacgggcaa 17880

cagctgattg ccccttcacc cctggccctg agagagttgc agcaagcggg ccacgctggt 17940

ttgcccagc aggcgaaaa cctgtttgat ggtggttcg aaatcggcaa aatccctat 18000

aaatcaaaag aatagcccgat gatagggttg agtgttggtc cagtttggaa caagagtcca 18060

ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 18120

ccactacgtg aaccatcacc caaatcaagt tttttgggt cgaggtgcgc taaagcacta 18180

aatcggaaac ctaaaaggag cccccgattt agagcttgac ggggaaagcc ggccgaacgtg 18240

gcgagaaagg aagggaagaa agogaaagga gggggcgcaa ttcaggctgc gcaactgttg 18300

ggaagggcga tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc 18360

tgcaaggcga ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac 18420

ggccagtgaa ttgagctcg gtaccggg 18449

<210> 41

<211> 18449

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 41

gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt 60

cctgtcaca ctaccaacat ggagtaocat aagggccagt tccgccagct cattaagagc 120

cagttcatgg gcgttggcat gatggcgcgc atgcacgtgt acttcaagta caccaacgct 180

cttctgatcc agtcgatcat ccgctgaagg cgttttcgaa tctggttaag atccacgtct 240

luggaayuc aygcactgyl yacctccagc gtccctttaa ggcctgcacac agctttctca 300

gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg 360
 gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg 420
 atattcttcg actctgtatt atagataaga tgataggaa ttggaggtag catagcttca 480
 tttggatttg ctttccaggc tgagactcta gcttggagca tagagggctc tttggctttc 540
 aatattctca agtatctcga gtttgaactt attccctgtg aacottttat tcaccaatga 600
 gcattggaat gaacatgaat ctgaggactg caatogccat gaggttttcg aaatacatcc 660
 gqatctcgaa gqctctggcc accctcgctg gttgaattta gaacgtggca ctattgatca 720
 tcgatagct ctgcaaggcg cgttgcacaa tgcaagtcga acgttgctag cagttccagg 780
 tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840
 caccacaaaa gtccagagcg gtaacacaaa gtcacacaa acagctgtga aggatttcgg 900
 caccggtacg gaagacggag aagccacctt cagtggaact gagtaccatt taattctatt 960
 tgtgtttgat cgagacctaa tacagccctt acaacacca tcaagtcgt atagctacca 1020
 gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa 1080
 ctatacagaa taagataggt ggagagctta taocgagctc ccaaatctgt ccagatoatg 1140
 gttgacoggt gcctggatct toctatagaa tcactcttat tcgttgacct agctgattct 1200
 ggagtgaacc agagggtcat gacttgagcc taaaatccgc cgcctccacc attttagtaa 1260
 aaatgtgaog aactcgtgag ctctgtacag tgaccoggtga ctctttctgg catgoggaga 1320
 gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380
 ctctgaggtg cagtgatga ttattaatcc gggaccggcc gccctccgc ccgaagtgg 1440
 aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcatcgagga atatggagct 1500
 tcactgaate accggcagta agcgaaggag aatgtgaagc caggggtgta tagcgtcgg 1560

cgaaatagca tgcattaac ctaggctacag aagtccaatt gcttcgcatc tggtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680

aattggccca tcgggcacat gtagggcgtc caaatatagt gcctctcctg ctttgcgcgg 1740

tgatatgaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtccctggta 1860

ggcagctttg ccccgctctg ccgcccgggt tgcggcgagg gttgacaagg tcgttgcgtc 1920

agtcacacat ttgttgccat attttctctg tctcccccac agctgctctt tttttttctc 1980

tttcttttcc catcttcagt atattcatct tcccatcaa gaacctttat tcccctaag 2040

taagtacttt gctacatcca tactccatcc tcccatccc ttattccttt gaacctttca 2100

gttcgagett tccacttca tcgcagcttg actaacagct acccgccttg agcagacatc 2160

accatgcctg aactcacgcg gacgtctgtc gagaagtttc tgatcgaaaa gtccgacagc 2220

gtctccgacc tgatgcagct ctccggagggc gaagaatctc gtgctttcag cttcgatgta 2280

ggagggcggt gatatgtcct gcgggtaaat agctgcgcgg atggtttcta caaagatcgt 2340

tatgtttatc ggcacatttg atogggcgcg ctcccgatgc cggaaagtgt tgacattggg 2400

gaattcagcg agagcctgac ctattgcato tcccgccgtg cacagggtgt cacgttgcaa 2460

gacctgcctg aaaccgaact gcccgctggt ctgcagccgg tcgaggaggc catgcatcgc 2520

atcgctgcgg ccgatatagg ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580

ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640

tggcaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700

atgctttggg ccgaggactg ccccgaagtc cggcacctcg tgcaocggga tttcggctcc 2760

aacaalgltc tgaaggacaa tggccgcata acagcggtca ttgactggag cggggcgatg 2820

ttcgggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880

atggagcagc agacgcgcta ctctgagcgg aggcacccgg agcttgacgg atccgcgcgg 2940

ctccggggcg atatgctcgg cattggctct gacaaactct atcagagctt ggttgacggc 3000

aatttcgatg atgcagcttg ggcgcagggg cgatgcgaag caatcgtcgg atccggagcc 3060

gggactgtcg ggcgtacaca aatogcccg agaaagcgcg cgtctgggac cgatggctgt 3120

gtagaagtac tcgcogatag tggaaaccca cgcgccagca ctgctccgag ggcaaggaa 3180

tagagttagt gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240

tgacgaatct ggaataaga tcgttgggtg cgatgtcagc tcggaggttg agacaaatgg 3300

tgttcaggat ctogataaga tacgttcatt tgtccaagca gcaaagagtg cttctagtg 3360

atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct ctccagata 3420

cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480

gcgaagagaa gaatagctta gcagagctat ttcatatttc gggagacgag atcaagcaga 3540

tcaacggctg tcaagagacc tacgagactg aggaatccgc tcttggtctc acgcgactat 3600

atatttgtct ctaattgtac ttgacatgc tctcttctct tactctgata gcttgactat 3660

gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tctctgaact 3720

ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan tctctactaa 3780

gatggtatac aatagtaacc atgcattggt gcctagttaa tgctccgtaa caaccaatac 3840

gcggcccgaa acttttttac aactctccta tgagtgtttt acccagaatg cacagggtaca 3900

cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgcoca 3960

ctccacatct ccactcgacc tgcaggcatg caaagcttga gattaaaata gataaggaaa 4020

agaaagttaa aagaaattcg gaagcatggc acattcttct tttataaat acatgcccta 4080

ctttcttttt ccctcgatat gatatatgca tatgatagat atacaagcaa tcttcttcaa 4140

ggagtttgaa attttgtcct ccaggagcaa aaaaagttt tttttalac atgtttgtac 4200

acaagaatag ttaccaattt gctttggtct taogtgotgc aagtttatat cgttttcaat 4260

ttctttgtct ttacatttct tttgtccttt atcttctctc atttagtctt tgggagaatt 4320

aggaaaaggg agcggaaagg taagaaatgc ttgcttattt tactaattcg gcaaacatcc 4380

aattttggca acagcagcct gtgcaacgct ctcgagatga cagtatcttt gattacactc 4440

taaatctcga tgaccgcacc aaaaagagcg aacaagaaa taatcttgtg cattcgaata 4500

tgatggaaga ttttttcccc cttattctaa atgttgacat agcgtgtatg ttatataaac 4560

aaaaagaaal lylacaaact ttcttttctt ctctttttat tttatctcta tgcgtctgaa 4620

gctgcagtca atcagcgtca agcccccgcg cgttgaacta gcccgcgaca tcacgcggcc 4680

caaagtctgc ctgcatgctc agcgggtgct gttagttcgg ctgcgagtgg cagcaccaca 4740

gacagaggag gcgctgggaa ccgtgcaggc tgccggcgcg ggcgatgagc acagcgccga 4800

tgtagcactc cagcagcttg accgggctat cgcagagcgt cgtgcccgcc gcaaacggga 4860

gcagctgtca taccaggctg ccgccattgc agcatcaatt ggcgtgtcag gcattgccat 4920

cttgcgccacc tacctgagat ttgccatgca catgacogtg ggcggcgagc tgccattggg 4980

tgaaagtggc ggcaactctc tcttggtggt tgggtggcgcg ctggcatgg agatgtatgc 5040

ccgctatgca cacaagcga tctggcatga gtgcctctg ggctggctgc tgcacaagag 5100

ccaccacaca cctgcactg gaacctttga agccaaogac ttgtttgcaa tcatcaatgg 5160

actgcccgcc atgtctctgt gtacctttgg cttctggctg cccaacgtcc tggggggcgc 5220

ctgctttgga ggggggctgg gcatacagct atacggcatg gcataatatgt ttgtacacga 5280

tggccttggtg cacaggcgct ttccacccgg gccatcgct ggctgccc acatgaagcg 5340

cctgacagtg gccaccagc tacaccacag cggcaagtac ggtggcgccg cctgggggat 5400

gttcttgggt ccacaggagc tgcagcacat tccagglycg gcggaggagg tggagcyac 5460

ggctcctggaa ctggactggt ccaagcgggc gattgtgact gatagcgaga ctctgggtcg 5520

atgttatctg cctcaacaat ggcttagaaa agaagaaca gaacaatac agcaaggcaa 5580

cgcccgtagc ctagggtgac aaagactggt gggctgtct ctgaagcttg taggaaaggc 5640

agacgctatc atgggtgagc ctaagaaggc cattgacaag ttgccggcaa actgtcaagg 5700

cggtgtacga gctgcttgc aagtatatgc tgcaattgga tctgtactca agoagcaga 5760

gacaacatat cctacaagag ctcatctaaa aggaagcgaa cgtgccaaga ttgctctggt 5820

gagtgtatac aacctctatc aatctgaaga caagcctgtg gctctccgtc aagctagaaa 5880

gattaagagt ttttttctg attagtgaat ttttcttta ttatgtctg atagtccaat 5940

aaagagacaa cacatacaat ataaaatcat tgtctttaa tggtaattta gttagtgtta 6000

aagcctgcat ttttttctg cgataaaca atgaattcac ccgctctcg gtttttaaat 6060

aattatgtca aactaggcaa aatctctttt ttctctctcg ttctttttt ggcttgttgt 6120

ggagtacag gctgtcttc agattgatag aggttgtata cactcaacag agcaatcttg 6180

gcacgttcgc ttcttcttag atgagctctt gtaggatatg ttgtctctg ctgcttgagt 6240

acagatccaa ttgcagcata tacttggcaa gcagctcgta cccgccttg acagtttgc 6300

ggcaacttgt caatgccctt cttagctctc accatgatag cgtctgcctt tctacaagg 6360

ttcagagaca agcccaacag tctttgatca cctaggctac gggcgttgc ttgctgtatt 6420

tgttctgttt ctctctttct aagccattgt tgggcagat aacatcgacc caacatctc 6480

gagccatact acagcataaa aggatacgtt ttctttaaca gaaatttacc cttttgttat 6540

cagcacatcac aaaaaaaaaag aaattttaaga tgagtagggac ttccattctc tcaaaaatttt 6600

tattcaatcc ataatgaat tatttttggg caaaaagaa agattatgcc tgattttctc 6660

tatttttttt tttlllcaaa ctcaccaat actttctagc ccagcttggc gtaatcatgy 6720

tcatactgt ttctgtgtg aaattgttat ccgctcaca ttccacacaa catacgagcc 6780

ggaagcataa agtgtaaac ctgggggtgcc taatgagtga gtaactcac attaatggcg 6840

ttgcgtcac tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 6900

ggccaacgcg cgggggagag cggtttgcgt attggggcaa agacaaaagg ggcacattca 6960

accgattgag ggagggaagg taaatattga cggaaattat tcaataaagg tgaattatca 7020

ccgtaccga cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta 7080

ccattagcaa ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta 7140

gcgacagaat caagtttgcc tttagcgtca gactgtagcg cgttttcacg ggcattttcg 7200

gtcatagccc ctttattagc gtttgccatc ttttcataat caaaatcacc ggaaccagag 7260

ccaccaccgg aaccgcctcc ctgagagcgg ccaccctcag aaccgccacc ctgagagcca 7320

ccaccctcag agccgccacc agaaccaaca ccagagccgc cgccagcatt gacaggaggg 7380

ccgatctagt aacatagatg acaccgcgcg cgataattta tctagttttg cgcgtatat 7440

tttgttttct atcgcgtatt aaatgtataa ttgcgggact ctaatcataa aaaccatct 7500

cataaataac gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa 7560

ttatatgata atcatcgcaa gaccggcaac aggattcaat ctaagaaac tttattgcca 7620

aatgtttgaa cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga 7680

acgcagcaag atatgcgggt gcactctcgt cttgcctggg cagtcgccgc cgacgcgctt 7740

gatgtgyaog ccgggcccga tcatattgtc gctcaggatc gtggcyltlt gullgclgyc 7800

cgttctgtgc gtaatgatat cggcacccttc gaccgcctgt tccgcagaga tcccgtaggc 7860

gaagaactcc agcatgatag ccccgcgctg gaggatcacc cagcggcggt ccggaaaaac 7920

gattccgaag cccaaccttt catagaaggg ggccgtggaa tcgaaatctc gtgatggcag 7980

gttgggctgc gcttggtcgg tcatctcgaa ccccgagtc ccgctcagaa gaactcgtca 8040

agaaggcgat agaaggcgat gcgctcgaa tcgggagcgg cgataccgta aagcacgagg 8100

aagcggctag ccattcgcgc gccaaagctct tcagcaatat cacgggtagc caacgctatg 8160

tcctgatagc gylcugucac acccagccgg ccacagtcga tgaatccaga aaaggggcca 8220

ttttccacca tgatatcgg caagcaggca tcgccatggg tcacagcgag atcatcgccg 8280

tcgggcctgc gcgccttgag cctggcgaa agtctggctg gcgcgagccc ctgatgctct 8340

tcgtccagat catcctgac gacaagacg gcttccatcc gagtacgtgc tcgctcgatg 8400

cgatgtttcg cttggtgtgc gaatgggcag gtacccgat caagcgatg cagccgcgc 8460

attgcctcag ccattgatga tactttctcg gcaggagcaa ggtgagatga caggagatcc 8520

tgccccggca cttcgcccaa tagcagccag tccttcccg cttcagtgac aacgtcgagc 8580

acagctgcgc aaggaaacgc cgtcgtggcc agccacgata gccguclyc utcgtcctgc 8640

agttcattca gggcacccga caggtcggtc ttgacaaaa gaacggggcg ccctgcgct 8700

gacagccgga acacggggc atcagagcag ccgattgtct gttgtgccc gtcataggcg 8760

aatgcctct ccacccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg 8820

cgaaacgac cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt 8880

ggataccgag gggaaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta 8940

gctgatagt accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt 9000

agtcattaa actccagaaa ccccgggctg agtggctcct tcaacgttgc ggtctctgta 9060

gttccaaacg taaaacggct tgtccgcgt catcgccggg ggtcataacg tgactccctt 9120

aattctccgc tcatgatcag attgtcggtt ccgccttca gtttaaacta tcaagtgttg 9180

acaggatata ttggcgggta aacctaaagc aaaagagcgt ttattagaat aatcggat 9240

ttaaaggcgt gtgaaaaggt ttatccgttc gtccatttgt atgtgcacgc caaccacagg 9300

gttccccaga tctggccgcg gccagcgaga cgagcaagat tggccgccgc ccgaacgat 9360

ccgacagcgc gccagcaca ggtgcgcagg caaattgcac caacgcatac agcgccagca 9420

gaatgccata gtggcggtg acgtcyltuy agtgaaccag atcgccagg aggcocggca 9480

gcacccgcat aatcaggccg atgcgcagc cgctcgagcgc gacagtgtc agaattacga 9540

tcaggggtat gttgggttc acgtctggcc tcgggaccag cctccgctgg tccgattgaa 9600

cgcgcggtt ctttatcact gataagttgg tggacataatt atgtttatca gtgataaagt 9660

gtcaagcatg acaaagttgc agcgaatac agtgatccgt gccgccttg acctgttgaa 9720

cgaggtcgc gtagacggtc tgagcacacg caaactggcg gaacggttgg gggttcagca 9780

gccggcgctt tactggcact tcaggaaaca gccggcgctg ctgcacgcac tggccgaagc 9840

catgctggcg gagaatcata cgcattcggg gccgagagcc gacgacgact ggcgtcact 9900

tctgatccgg aatgcccgca gcttcaggca ggcgctgtc gctaccgcg atggcgcgcg 9960

catccatgcc ggcacgcgac cgggcgcacc gcagatggaa acggccgacg cgcagcttcg 10020

cttcctctgc' gaggcgggtt ttccggccgg ggacgcgctc aatgcgctga tgacaatcag 10080

ctacttact gttggggcgg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg 10140

cgcgccgacc gttgaacagg ctccgctctc gccgctgtg cgggcgcgca tagacgctt 10200

cgacgaagcc ggtccggcag cagcgttcga gcagggaact gcggtgattg tcatggatt 10260

ggcgaagg aggcctgtg tcaggaaact tgaaggaccg agaagggtg acgattgatc 10320

aggaccgctg ccggagcgca acccactcac tacagcagag ccatgtagac aacatccctc 10380

ccccctttcc accgcgtcag acgcccgtag cagcccgtc cggtcttttt catgccctgc 10440

cctagcgtec aagcctcac gccgcgctcg gcctctctgg cggccttctg gcgctcttcc 10500

gcttctctcg tcactgactc gctgcgctcg gtcttctggc tgcggcgagc ggtatcagct 10560

cactcaaaag cggtataacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 10620

tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 10680

cataggtctc gccccctga cgaacatcac aaaaatcgac gctcaagtca gaggtggcga 10740

aaccgcacag gactataaag ataccaggcg ttccccctg gaagctccct cgtgcgctct 10800

cctgttccga ccctgccgct tacccgatac ctgtccgctt ttctcccttc ggggaagcgtg 10860

gcgcttttcc gctgcataac cctgcttcgg ggtcattata cgcatttttt cggtatatcc 10920

atccttttcc gcaacatata caggattttg ccaaaggggt cgtgtagact ttcccttggtg 10980

tatccaacgg cgtcagccgg gcaggatagg tgaagtaggc ccaccgcga gcgggtgttc 11040

ctttctcact gtcccttatt cgcacctggc ggtgctcaac gggaaacctg ctctgcgagg 11100

ctggccggct accgcggcgg taacagatga gggcaagcgg atggctgatg aaaccaagcc 11160

aaccggaag ggcagcccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat 11220

tgaggaaaag gcggcggcgg ccggcatgag cctgtcggcc tacctgctgg ccgtcgcca 11280

gggtacaaa atcacggcgg tcgtggacta tgagcacgtc cgcgagctgg ccgcacataa 11340

tggcgcacct ggcgcgctgg gcggcctgct gaaactctgg ctccacgacg acccgcgcac 11400

ggcgcggttc ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga 11460

cgagcttggc aaggtcatga tgggcgtggt ccgcccgagg gcagagccat gactttttta 11520

gcccgtaaaa cggccggggg gtgcgcgtga ttgccaagca cgtcccatg cgtccatca 11580

agaagagcga ctgcgcggag ctggtgaagt acatcaccca cagcaaggc aagaccgagc 11640

gcctttgoga cgtccaccgg gctggttgcc ctgcgcgtg ggtgggggc cgtctatggc 11700

cctgcaaacy cgcagaaaaa gccgtcgaag cgtgtgcga gacaccggc cgcggggcgt 11760

tgtggatacc tcgcggaaaa ctggccctc actgacagat gaggggcgga cgttgacact 11820

tgaggggcgc actcaccggc cgcggcggtg acagatgagg ggcaggctcg atttcggcgc 11880

gcgacgtgga gctggccagc ctgcgaaatc ggcgaaaaac cctgatttta cgcgagtttc 11940

ccacagatga tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc 12000

gcgactactg acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga 12060

tgaggggcgc acctattgac atttgagggg ctgtccacag gcgaaaaac cagcatttgc 12120

aaggggttcc gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaacca 12180

atatttataa acctgtttt taaccagggc tgccccctgt gcgcgtgacc gcgcacgcgc 12240

aaggggggtg ccccccttc tcgaaccctc cgggccgcct aacgcgggcc tcccatcccc 12300

ccaggggcgt cgcctctcgg ccgcgaacgg cctcacccca aaaatggcag cgtgggcagt 12360

ccttgccatt gccgggatgc gggcagtaac gggatgggcg atcagccga gcgcgacgc 12420

cggaaacatt gacgtgcgcg aggtgctggc atcgacattc agcgaccagg tgcggggcag 12480

tgaggggcgc ggcctgggtg gcggcctgcc ctcaacttc gcgtcgggg cattcacgga 12540

cttcattggc gggccggcaa tttttacctt gggcattctt ggcattagtgg tcgggggtgc 12600

cgtgctcgtg ttccgggggtg cgataaacc agcgaaacct ttgaggtgat aggtaaagatt 12660

ataccgaggt atgaaaacga gaattggacc ttacagaat tactctatga agcccatat 12720

ttaaaagct accaagacga agaggatgaa gaggatgagg aggcagattg cttgaaat 12780

attgacaata ctgataagat aatatactt ttatatagaa gatatcgccg tatgtaagga 12840
tttcaggggg caaggcatag gcagcgcgct tatcaatata tctatagaat gggcaaagca 12900
taaaaacttg catggactaa tgcttgaaac ccaggacaat aaccttatag ctgttaaatt 12960
ctatcataat tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgcc 13020
gatgactttg tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtccagccg 13080
tgccagggtc tgccctcagat tcagggttatg ccgctcaatt cgctgcgtat atcgcttgct 13140
gattactgtc agctttccct tcaggcggga ttcatacagc gccagccat ccgtcatcca 13200
tatcaccacg tcaaaggggt acagcaggct cataagacgc ccagcgctgc ccatagtgcg 13260
ttcacccaat acgtgcgcaa caaccgtctt ccgggagctg tcatacgctg aaaacagcca 13320
gcgctggcgc gatttagccc cgacatagcc ccactgttcg tccatttcgc cgcagacgat 13380
gacgtcactg cccggctgta tgcgcgaggt taccgactgc ggcctgagtt ttttaagtga 13440
cgtaaaatcg tgttgaggcc aacgcccata atcggggctg ttgcccgga tccaacgcca 13500
ttcatggcca tatcaatgat tttctggtc gtaccgggtt gagaagcgtt gtaagtgaac 13560
tgcagttgcc atgttttacg gcagtgcgag cagagatagc gctgatgtcc ggcggtgctt 13620
ttgcggttac gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa 13680
gccactggag cacctcaaaa acaccatcat acactaaac agtaagttgg cagcatcacc 13740
cataattgtg gtttcaaaat cggctccgct gatactatgt tatacgccaa ctttgaaac 13800
aactttgaaa aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg 13860
gagttcgtct tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa 13920
ggaaataata aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat 13980
accgctgcgt aaaagatagc gaaggaaatgt ctctgctaa ggtatataag ctggtgggag 14040

aaatgaaaa cctatattta aaaatgacgg acagccggta taaagggacc acctatgatg 14100

tggaacggga aaaggacatg atgctatggc tggaaggaaa gctgcctggt ccaaagggtc 14160

tgcactttga acggcatgat ggctggagca atctgctcat gagtggagcc gatggcgctc 14220

tttgcctgga agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg 14280

aqtgcacag gctcttttca tccatcgaca tatcggattg tccctatacg aatagcttag 14340

acagccgctt agccgaattg gattacttac tgaataacga tctggcgatg gtggattgcg 14400

aaaactggga agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga 14460

cgaaaaagcc cgaagaggaa cttgtctttt ccacgcgcga cctgggagac agcaacatct 14520

ttgtgaaaga tggcaaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca 14580

agtggtatga cattgccttc tgcgtccggt cgatcagggg ggatatacgg gaagaacagt 14640

atgtcgagct attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaaatatt 14700

atatthttact ggatgaattg ttttaqtacc tagatgtggc gcaacgatgc cggcgacaag 14760

caggagcgca ccgacttctt ccgcatcaag tgttttggct ctacggccga ggcccacggc 14820

aagtatttgg gcaaggggtc gctggtatto gtgcagggca agattcggaa taccagtac 14880

gagaaggacg gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg 14940

gacaccaagg caccagcggt gtcaaatcag gaataagggc acattgcccc ggctgagtc 15000

ggggcaatcc cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa 15060

gaactgatog acgcggggtt ttccgcgcag gatgcccga ccatcgcaag ccgcaccgtc 15120

atgogtgcgc ccgcgaaac cttccagtc ctcggctoga tggctcagca agctacggcc 15180

aagatcgagc gcgacagcgt gcaactgggt cccctgccc tgcgcgcgc atcgccgcgc 15240

gtggagcgtt cgcgtcgtct cgaacaggag gcggcaggtt tggcgaagtc gatgaccac 15300

gacacgcgag gaactatgac gaccaagaag cgaaaaacgg ccggcgagga cctggcaaaa 15360

caggtcagcg aggccaaagca ggcgcggttg ctgaaacaca cgaagcagca gatcaaggaa 15420

atgcagcttt ccttgttoga tattgcgcgg tggccggaca cgatgcgagc gatgcaaac 15480

gacacggccc gctctgcctt gttcaccacg cgcaacaaga aaatcccgcg cgagcgcgctg 15540

caaaacaagg tcattttcca cgtoaacaag gacgtgaaga tcacctacac cggcgctcgag 15600

ctgcggggcg acgatgacga actgggtgtg cagcaggtgt tggagtaacg gaagcgcacc 15660

cctatcggcg agccgataac cttcacgttc tacgagcttt gccaggacct gggctggtcg 15720

atcaatggcc ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg 15780

atgggcttca cgtccgaacg cgttgggcac ctggaatcgg tgtcgctgct gcaccgcttc 15840

cgcgtcctgg accgtggcaa gaaaacgctc cgttgcacgg tctgatcgca cgaggaaatc 15900

gtcgtgctgt ttgctggcga coactacacg aaattcatat gggagaagta ccgcaagctg 15960

tgcgcgacgg ccgcacggat gttcgactat ttacgctcgc accgggagcc gtaaccgctc 16020

aagctggaaa ccttcgcctt catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc 16080

gagcaggctg gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctg 16140

gtcaatgatg acctggtgca ttgcaaacgc tagggccttg tggggtcagt tcggctggg 16200

ggttcagcag ccagcgcctt actggcattt caggaaacaag cgggcactgc tcgacgcact 16260

tgcttcgctc agtatcgctc gggacgcagc gcgcgcctca cgaactgcg ataacagag 16320

gattaaaatt gacaattgtg attaaggctc agattcgacg gcttgagcgc gccgacgtgc 16380

aggatttcgc cgagatccga ttgtggccc tgaagaaagc tcagagatg ttccgggtccg 16440

tttacgagca cgaggagaaa aagcccattg aggcgttcgc tgaacggttg cgagatgccg 16500

tggcattcgg cgcctacacg gacggcgaga tcattgggtt gtcggtcttc aaacagagag 16560

acggcccca ggcgcgtcac aaggcgcatc tgtcggcgt tttcgtggag cccgaacagc 16620
gaggccgagg ggtcgccggt atgctgctgc gggcggtgcc ggcgggttta ttgctcgtga 16680
tgatcgtccg acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac 16740
ttaatatctt gctattctgg agcttggtgt ttatttcggt ctaccgcctg ccggggcggg 16800
tcggggcgac ggtaggcgct gtgcagccgc tgatggctgt gtccatctct gccgctctgc 16860
taggtagccc gatacgattg atggcggtcc tgggggctat ttgcggaact gcggggcggtg 16920
cgctgttggt gttgacacca aacycagcgc tagatcctgt cggcgctgca gcgggcctgg 16980
cgggggcggt ttccatggcg ttcggaaccg tgctgaccgc caagtggcaa cctcccgctc 17040
ctctgctcac ctctaccgcc tggcaactgg cggcgggagg acttctgctc gttccagtag 17100
ctttagtgtt tgatcggcca atcccgatgc ctacaggaa caatgttctc ggctcgcgct 17160
ggctcggcct gatcggagcg ggtttaacct acttcccttg gttccggggg atctcgcgac 17220
tcgaacctac agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc 17280
cgggggatga tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcggtgag 17340
caatggatag gggagttgat atcgtaaacg ttcaactcta aayaaatagc gccactcagc 17400
ttcctcagcg gctttatcca gcgatttctt attatgtcgg catagttctc aagatcgaca 17460
gcctgtcacg gttaaagcag aaatgaataa gaaggctgat aattcggatc tctgcgaggg 17520
agatgatatt tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctacctt 17580
ccgcgagatc atcgtgtttt caaacccggc agcttagttg ccgttcttcc gaatagcacc 17640
ggttaacatg gcaaagctcg ccgccttaca aoggctctcc cgctgacgcc gtcgggactt 17700
gatgggctgc ctgtatcgag tgggtatttt gtgcogagct gccggctcgg gagctgttgg 17760
ctggctggtg gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac 17820

acattgcgga cgtttttaaat gtactggggt ggtttttctt ttcaccagtg agacgggcaa 17880
cagctgattg ccccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggt 17940
ttgccccagc aggcgaaaaat cctgtttgat ggtggttccg aaatcgggcaa aatcccttat 18000
aaatcaaaag aatagcccga gatagggttg agtggtgttc cagtttgaa caagagtcca 18060
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 18120
ccactacgtg aaccatcac caaatoaagt ttttggggg cgaggtgccg taaagcacta 18180
aatcggaacc ctaaaggagg ccccgattt agagcttgac ggggaaagcc ggcgaacgtg 18240
gcgagaaagg aagggaagaa agcgaaggga gcggggcgcca ttcaggctgc gcaactgttg 18300
ggaagggcga tcggtgcggg cctcttcgtt attacgccag ctggcgaaag ggggatgtgc 18360
tgcaaggcga ttaagttggg taacgccagg gttttccag tcadgacgtt gtaaaacgac 18420
ggccagtga ttagagctcg gtaccggg 18449

<210> 42

<211> 17593

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 42

```

ccggcgctggt tgccctcgcc gctggggtgg cggccgtcta tggccctgca aacgcgccag      60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtga tacctcgccg      120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac      180
ccggcgccgc gttgacagat ggggggcagg ctcgatllyc gccgcgcagc tggagctggc      240
cagcctcgca aatcggcgaa aacgcctgat tttaacgcgag ttccaccag atgatgtgga      300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat      360
gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat      420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggg ttccgccgt      480
ttttcgccca ccgctaaact gtcttttaac ctgcttttaa accaatattt ataaaccttg      540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc      600
cttctcgaac cctcccgccc cgctaacgcg ggctcccat ccccccaggg gctgcgcccc      660
tcggcgccga acggcctcac cccaaaaatg gcagcgctgg cagtccttgc cattgccggg      720
atcggggcag taacgggatg ggcgatcagc ccgagcgga cgcocggaag cattgacgtg      780
ccgcaggtgc tggcatcgac attcagcgac caggtgcggg gcagtgaggg cggcgccctg      840
ggtggcgccc tgcccttcac ttccggcgtc ggggcattca cggacttcac gggggggcgg      900
gcaattttta ccttgggcat tottggcata gtygtcgagg gtgcogtgc cgtgttcggg      960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa      1020
acgagaattg gacotttaca gaattactct atgaagcgcc atatttaaaa agctaccaag      1080

```

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140
agataatata tcttttatata agaagatata gccgtatgta aggatttcag ggggcaaggc 1200
ataggcagcg cgttatcaa tatatctata gaatgggcaa agcataaaaa cttgcattgga 1260
ctaagtcttg aaaccaggga caataacctt atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac ttgtcatgac 1380
agctccacgg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440
agattcaggt tatgcgcgc aattcgctgc gtatatcgct tgcgtattac gtgcagcttt 1500
cccttcaggc gggattcata cagcggccag ccattcgctc tccatatcac cactgcaag 1560
ggtagacga ggctcataag acgccccagc gtgcgcattg tgcgttcacc gaatacgtgc 1620
gcaacaacgg tcttcggag actgtcatc gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccacat agccccactg ttgcgtcatt tccgcgcaga cgatgacgct actgccccgc 1740
tgtatgcgcg aggttacga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatcggg gctgttgccc ggcattccaa gccattcatg gccatatcaa 1860
tgattttctg gtgcgtacgg ggttgagaag cgggtgaagt gaactgcagt tgcattgttt 1920
tacggcagtg agagcagaga tagcgtgat gtccggcggg gcttttgcgg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100
aaatcggctc cgtcgatact atgttatcgc ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggtt ttttaagggtt tagaatgcaa ggaacagtga attggagctc gtctgttat 2220
aattagcttc ttgggttctc ttttaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcacggg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaaggtata taagctggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggtcttt 2640

tcactccatc gacatatacg attgtcccta tacgaataga ttagacagcc gcttagccga 2700

attggattac ttaactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agccccgaaga 2820

ggaacttgtc ttttccacg gcgacctggg agacagcaac atcttttgta aagatggcaa 2880

agtaagtggc ttatttgatc ttgggagaag cggcagggcg gacaagtggc atgacattgc 2940

cttctgogtc oggtcgatca gggagatat cggggaaga cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaocgact 3120

tottocgat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggctcgtggt attcgtgcag ggcaagatc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacg gaccgactc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaaata gggcacattg ccccgcggtg agtcggggca atcccgcaag 3360

gaggytgaat gaatcggaag tttagccgga aggcatacag gaaagaaotg atogagoggy 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtcogtggc togatggtcc agcaagctac ggccaagatc gagcgcgaca 3540

gggtgcaact gggtccccc ggcctgcccg cgcacatcggc cgcctggag cgttcogctc 3600
 gtctogaaca ggaggcgcca ggtttggcga agtcgatgac catcgacacg cgagggaacta 3660
 tgacgaccaa gaagcgaaaa accgcgcgcg aggcactggc aaaaacagtc agcaggcgca 3720
 agcaggcgcg gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccctgt 3780
 tgcataattgc gccgtgcgcg cacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840
 cctgttcac cgcgcgcaac aagaaaaatcc cgcgcgaggc gctgcaaac aaggtcattt 3900
 tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgcgctgcgg gccgacgatg 3960
 acgaactggt gtgcgacgag gtgttgagat acgcgaagcg caccctatc ggcgagccga 4020
 tcacottcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggattt 4080
 aacgaaggc cgaggaatgc ctgtcgcgc caccggcgc ggcatgggc tccacgtccg 4140
 accgcgttgg gcaacctgaa tcggtgtcgc tgcgcacgc cttccgcgtc ctggaccgtg 4200
 gcaagaaaaa gtcocgttc caggtcctga tcgacagga aatcctcgtc ctgtttgctg 4260
 ggcaccacta cagcaaatc atatgggaga agtaccgca gctgtgcgc accgcccgc 4320
 ggatgttga ctatttcagc tcgcaccggg agcgttacc gctcaagctg gaacacctc 4380
 gctcatgtg cgcacggat tccaccgcg tgaagaagt gcgcgacgc gtcggcgaa 4440
 cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctg 4500
 tgcattgcaa acgctagggc cttgtgggt cagttccggc tgggggttca gcagcagcg 4560
 ctttactggc atttcaggaa caagcgggca ctgctcagc cacttgcttc gctcagtac 4620
 gtcgggagc caagcgcgcc tcacgaact gcgataaac agaggattaa aattgacaat 4680
 tgtgattaag gctcagattc gacggcttgg agcggccgc gtcagagatt tccgcgagat 4740
 cgcattgtgc gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860
catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980
cgggatgctg ctgcgggcgt tgcggcggg tttattgctc gtgatgatcg tccgacagat 5040
tccaaacgga atctcgtgga tgcgcattct catctcggc gcaactaata ttctcgtatt 5100
ctggagcttg ttgtttattt cggctacacg cctgcgggc ggggtcgagg cgaocgtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgcgct ctgctaggta gcccgatacg 5220
attgatggcg gtctcggggg ctatttgcgg aactgcgggc gtggcgctgt tgggtttgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga ccgcgaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgcttgccaa ctggcgcccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtgcctcg gccctgatcg 5520
agcgggttta acctacttcc tttggttcg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggctttctca gccccagatc tggggtcgat cagcggggga tgcatoaggc 5640
cgacagtccg aacttogggt cccgcacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttccat totaaagaa tagcgccact cagcttctc agcggcttta 5760
tcacgcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cagcggttaag 5820
cgagaaatga ataagaagcg tgataattcg gatctctcgc agggagatga tattttgatca 5880
caggcagcaa cgcctctgta togttacaat caacatgcta cctcccgga gatcatcgt 5940
gtttcaaac ccgcagctta gttgccgttc ttcgaatag catcggtaac atgagcaag 6000
tctgcgcct tacaacggt ctcccgctga cgcgctccg gactgatggg ctgcctgtat 6060

cgagtggtga tttttgtccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtggttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttggcc cagcaggcga 6300
aaatcctggt tgatggtggt tccgaatcg gcaaatccc ttataaatca aagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actocaaagt caaagggcga aaaaccgtct atcagggcga tggccacta cgtgaacct 6480
caccocaaatc aagttttttg gggtcgaggt gccgtaaagc actaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaagga 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcact gttgggaagg gcgatcggtg 6660
cgggcctctt cgtctattaog ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
tggttaacgc cagggttttc ccagtcacga cgttgtaaaa cgcggccag tgaattcgag 6780
ctcggtaacc ggggatcttt cgcactgaa atacgtcgag cctgctccgc ttggaagcgg 6840
cgaggagcct cgtcctgtca caactaccaa catggagtao gataaggyuc agttcccca 6900
gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtaactcaa 6960
gtacaccaac gctcttctga tccagtcgat catccgtga aggcgcttic gaattcgttt 7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080
aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgcccag 7140
aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200
cgggtgtcaag aggataattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctgcttgga gcatagaggy 7320

tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcggttgc caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagucac ulicaylyga ulicaylacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggaactc aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttatacagag ctcccaaatc 7920

tgtccagatc atggttgacc ggtgcctgga tcttctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga ccagagggt catgacttga gctctaaatc cgcgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaaccg tgactctttc 8100

tggcatgogg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgccctc 8220

cgcgccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcattcatcg 8280

agaatatgga gcttcacga atcacccgca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aaactaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgtgtatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cgggaaacaca agctggcagt cgaccocatcc ggtgctctgc actcgacctg ctgagggtccc 8640

tcagtcacctg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700

aggctgttgc gtcagtcocaa catttgttgc catattttcc tgcctctccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttccctt aagtaagtac ttgtctacat ccatactcca tcttcccat cccttattcc 8880

tttgaacott tcagttcgag ctttcccaact tcctgcagc ttgactaaca gctaccocgc 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatoga 9000

aaagtgcag agcgtctccg acctgatgca gctctcggag ggogaagaat ctcgtgcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatgggtt 9120

ctacaaagat cgttatgttt atcgccactt tgcctcggcc gcgctccoga ttcgggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctccgcgc gtgcacaggg 9240

tgctacgttg caagaacctg ctgaaacoga actgcocgct gttctgcagc cggctgcgga 9300

ggccatggat gcgctcgtc cggccgatct tagccagacg agcgggttcg gcccatcgg 9360

accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacacgcgc agtgcgtccg tcgcgcaggg 9480

tctgatgag ctgatgcttt gggccgagga ctgcoccgaa gtccggccacc tcgtgcacgc 9540

ggatttcgcg tccaacaaat tctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgagggc atgttcgggg attccaata caggtgcgcc aacatcttct tctggaggcc 9660

gtggttggtc tgtatggagc agcagacgcg ctacttcgag cggaggccatc cggagcttgc 9720

aggatgcgcg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag gctcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtag acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac gcacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgcgcacc gcgggatcga tccacttaac gtactgaaa 10020

tcatcaaaac gcttgacgaa tctggatata agatcgttg tgcgatgtc agctcggag 10080

ttgagacaaa tgggtgtcag gatctcgata agatacgttc atttgtcaa gcagcaaga 10140

gtgccttcta gtgatttaat agctccatgt caacaagaat aaacgcggtt ttccgggtta 10200

cctcttcag atacagctca tctgcaatgc attaatgca lyaclgcaac ctactgacgc 10260

cttnnaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttccggagac 10320

gagatcaagc agatcaacgc tcgtcaagag acctacgaga ctgaggaatc cgtctctggc 10380

tccacgcgac tatatatctg tctctaattg tactttgaca tgcctctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtaccag cncctgggtt cgcgaagata attgcatgtt 10500

tcttccttga actctcaagc ctacaggaca cacattcacc gtagggtata acctcgaaat 10560

canttcctac taagatggta tacaatagta accatgcacg gttgcctagt gaatgctccg 10620

taacacccaa tagccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680

atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctctgttact 10740

gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt ttttcgagtt 10800

tttttttttt ttctttgtga aggatttatt gttattggta tccatttttt attggaagac 10860

aagatacgtt aatattgatt ttgcttaag attaaaggga aatcagaaaa cgacataaa 10920

aatgtzaag gacaaactat ggtgtcgatt ataagtctaa atccttaaaa aatgacaacg 10980

agttgcttct ctctgaaac aattcttttg tctttgcaag aaagggttct tttttgtttg 11040

cttgcattac ttaaacaatca aatcaaatga aaggaataaa gcagatttga gggcgaataa 11100

ggattttctg gtcaacaaga tgtgagtgac acctaaggaa ctaaagtcca ttcatttggt 11160

ttaaacgcac atcaagatt gatgatcaac aggattgaga gagagaaaaa gaactcgtgt 11220

cattttattc tgttgactga aattttatat ttgaaaaaa tgccaatct atagctttag 11280

ctatattaca taacatttga aataataata ataaaaaaag acacattaga gacacttttc 11340

aaactctaaa taactgtcta taaacacaaa gaaaacaaag acctctataa caacttatta 11400

gatttttctc gtacttttgt ctaaagatga tgtattcttg ttatcccaca ctcttttcat 11460

ttgtttctga tgtactataa tatacaaaat ttcttttttg caagagatat tattccaaaa 11520

atttcaaaa agaaatttt ttcacaatag cagttgatcg tgtaacccaa agaggttctt 11580

tgttattttg cacttcgcgt ttgcggtgat gcattattcaa agtaatatat ggaataaaca 11640

acgtgtttaa gcatgaaga aaggaaacaa aggcgcgttt gaacaaatgc ataattttc 11700

agcaaaaaat gatctaaagc aacagtaaa tcaacaaga aacattgctg attcgcgtta 11760

gaaaacgata aagtccta atagccactaa gtatacttca atgaactttt tgtatgctta 11820

tggtccaato agaccaataa ttgtgacca ttcttgaggt ggcttttggt atgcggaac 11880

agaaaaaaat tttctacca atcgatttaa aaacaattt ctgctttgaa ccaaaacttt 11940

ttttttctct ttaatcatta actttatcaa gtatgtacct accctcaaag tctcactca 12000

agcacatta tgcatacatt gttccacct ctctttagaa atgttgtgga ttggaatgc 12060

cctgatcgtt ttcgttacg tgattggcat ggaagtgatt gctgcactgg cacacaaata 12120

cactcatgac ggtgggggt ggggatggga tctttcacat catgaaccgc gtaaaagtgc 12180

gtttgaagtt aacgatctt atgcogtgg ttttgtgca ttatgatcc tgttgattta 12240

tctgggcagt acaggaatgt ggccgctcca gtggattggc gcaggtatga cggcgatgg 12300

attactctat tttatggtgc acgacgggct ggtgcatcaa cgttggccat tcgcctatat 12360
tccacgcagc ggctacctca aacgggttga tatggcgcac cgtatgcac acgcgcgcag 12420
gggcaaaaga ggttggtgtt cttttggctt cctctatgcg ccgccctgt caaaacttca 12480
ggcgacgcgc cgggaaagac atggcgctag agcggggcgt gccagagatg cgcaggcgcg 12540
ggagatqag ccgcacccg ggaagtaagg gcctgaccag agggggccag cagcagcggt 12600
aatttttcgg gcgtggtcgt tgactgccgc tgatcccaaa gcttggcgta atcatgggtca 12660
tagctgtttc ctgtgtgaaa ttgttatccg ctcaaatc cacaacat acgagccgga 12720
agcataaagt gtaaagcctg gggcgctaa tgagtgcgt aactcacatt aattgcgttg 12780
cgctcactgc ccgctttcca gtggggaac ctgtcgtgcc agctgcatta atgaatcggc 12840
caacgcgcgg ggagaggcgg tttcgtatt gggccaaaga caaaggcg acattcaacc 12900
gattgagga ggaaggtaa atattgacg aaattattca ttaagggtga attatcacgg 12960
tcaccgactt gagccattt ggaattagag ccagcaaat caccagtagc accattacca 13020
ttagcaaggc cggaaacgct accaatgaaa ccacgatag cagcacgta atcagtagcg 13080
acagaatcaa gtttgcttt agcgtcagac tgtagcgggt ttcatcggc attttcggtc 13140
atagccccc tattagcgtt tgccatcttt tcataatcaa aatcacgga accagagcca 13200
ccaccggaac cgctccctc agagccgcca ccctcagaac cgcccccctc agagccacca 13260
ccctcagagc cgccaccaga accaccacca gagccgcgc cagcattgac aggaggcccg 13320
atctagtaac atagatgaca ccgcgcgcga taatttatcc tagtttgcgc gctatatatt 13380
gttttctatc gcgtattaaa tgtataattg ogggactcta atcataaaaa cccatctcat 13440
aaataacgct atgcattaca tgtaattat tacatgctta acgtaattca acagaaatta 13500
tatgataatc atcgcaagac cggcaacagg attcaatctt aagaaacttt attgcaaat 13560

gtttgaacga tcggggatca tcggggtctg tggcggggaac tcacgaaaa tatccgaacg 13620

cagcaagata tcggggtgca tctcggtctt gcctgggcag tcggcgccga cgcggttgat 13680

gtggagcccg ggcgcgatca tattgtcgct caggatcggt gcgttggtct tgtgggcctt 13740

tgctgtcgta atgatatcgg caccttcgac gcctgttcc gcagagatcc cgtggggcaa 13800

gaactccagc atgagatccc gcgctggag gatcatccag ccggcgctcc ggaanaacgat 13860

tcgaagccc aacctttcat agaaggcggc ggtggaatcg aaatctcgtg atggcaggtt 13920

gggcgtcgct tggtcggtca ttctgaaccc cagagtccc ctcagaagaa ctogtcaaga 13980

aggcgataga aggcgatgag ctgcgaatcg ggagcggcga tacgtataag caccaggaag 14040

cggtcagccc attcgccgc aagctcttca gcaatatac gggtagccaa cgctatgtcc 14100

tgatagcggc cgcgcaccc cagccggcca cagtogatga atccagaaaa cgggccattt 14160

tocaccatga tattcgga cagggcatcg ccattgggtca cgaagagatc atgcgcgtcg 14220

ggcatcgccg ccttgagcct ggcgaacagt tcggctggcg cgaacccctg atgctcttcg 14280

tcagatcat cctgatcgac aagaccggct tccatccgag taactgctcg ctogatgga 14340

tgtttcgctt ggtgggtcaa tgggcaggta gcggatcaa gcgtatcgac ccgcgcgatt 14400

gcacagcca tgatggatc ttctcggca ggagcaaggc gagatgacag gagatcctgc 14460

ccggcactt cgcgaatag cagccagtc ctcccgctt cagtacaac gtgcgacaa 14520

gctgcgcaag gaacgccgt cgtggccagc cagatagcc gcgctgcctc gtctcgagc 14580

tcattcaggc caccggacag gtccgtcttg aaaaaagaa ccgggcgccc ctgcgctgac 14640

agccggaaca cggcgcatc agagcagcc attgtctgtt gtgccagtc atagcgaat 14700

agcctctcca ccaagcggc cggagaacct gcgtcaatc catcttgctt aatcatgga 14760

aacgatccag atccggtgca gattatttgg attgagagtg aatatgagac tctaattgga 14820

taccgagggg aatttatgga acgtcagtg agcatttttg acaagaaata ttgtctagct 14880
gatagtgaacc tttaggcgact: ttggaacgcg caataatggt ttctgacgta tgtgcttagc 14940
tcattaaact ccagaaaccc gcggtcagtg ggctccttca acgttgoggt tctgtcagtt 15000
ccaaacgtaa aacggcttgt cccgcgtcat cggcggggggt cataacgtga ctcccttaat 15060
tctccgctca tgatcagatt gtogtttccc gccttcagtt taaactatca gtgtttgaca 15120
ggatatattg gcgggtaaac ctaagagaaa agagcgttta ttagaataat cggatattta 15180
aaagggcggtg aaaagggtta tccyllcgtc catttgtatg tgcattgccaa ccacaggggt 15240
ccccagatct ggccgcccgc agcgagacga gcaagattgg ccgcgcgcgc aaacgatccg 15300
acagcgcgoc cagcacaggt gcgcaggcaa attgcaccaa cgcatacagc gccacgagaa 15360
tgccatagtg ggcggtgacg tcgttcgagt gaaccagatc gcgcaggagg ccgcgcgaca 15420
ccgcgataat caggccgatg ccgacagcgt cgagcgcgac agtgctcaga attacgatca 15480
ggggtatggt gggtttcaag tctggcctcc ggaccagcct ccgctgggtcc gattgaacgc 15540
gcggattcct tctactgat aagttggtgg acatattatg tttatcagtg ataaagtgtc 15600
aagcatgaca aagttgcagc cgaatacagt gatccgtgcc gccctggacc tgttgaaacya 15660
ggtcggcgta gacggtctga cgacacgcaa actggcgga cggttggggg ttcagcagcc 15720
ggcgctttac tggcacttca ggaacaagcg ggcgctgctc gacgcactgg ccgaagccat 15780
gctggcgag aatcatacgc attcggtgcc gagagccgac gacgactggc gctcatttct 15840
gatcggaat gcccgagct tcaggcagcg gctgctcgcc tacgcgatg gcgcgcgcat 15900
ccatgcggcg acgcgacggg gcgcaccgca gatggaaacg gccgacgcgc agcttogett 15960
cctctcgag gcgggttttt cggccgggga cgcctcaat gogctgatga caatcagcta 16020
cttactggt ggggcggtgc ttgaggagca ggccggcgac agcgatgcgc gcgagcgcgg 16080

cggcaccggtt gaacaggctc cgctctcgcc gctgttcggg gcgcgatatg acgccttcga 16140

cgaagccggt cccgacgcag cgttcgagca gggactcgcg gtgattgtcg atggattggc 16200

gaaaaggagg ctgctgtgca ggaacgttga aggaccgaga aagggtgacg attgatcagg 16260

accgctgccg gagcgcaacc cactcactac agcagagcca ttagacaac atccccccc 16320

cttttcacc gcgtcagacg cccgtagcag cccgctacgg gctttttcat gccctgccct 16380

agcgtccaaag cctcacggcc gcgctcgcc tctctggcgg ccttctggcg ctcttccgct 16440

tcctcgctca ctgactcgct gcyctcggtc gttcggctgc ggcgaggggt atcagctcac 16500

tcaaaggcgg taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga 16560

gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggcgg cgttgctggc gtttttccat 16620

aggctccgcc cccctgacga gcatacaaaa aatcgacgct caagtacagag gtggcgaaaac 16680

ccgacaggac tataagata ccaggcggtt cccctggaa gctccctcgt gcgctctcct 16740

gttcgcacc tcgcgcttac cggatacctg tccgccttcc tcccttcggg aagcgtggcg 16800

cttttcgct gcataaccct gcttcggggg cattatagcg attttttcgg tatatccatc 16860

ctttttcgca cgatatacag gattttgcc aaggggtcgt gtagacttcc ctgggtgtat 16920

ccaacggcgt cagccgggca ggatagggtga agtagggcca ccgcgagcg ggtgttcctt 16980

cttccactgc ccttattcgc accctggcgg gctcaacggg aatcctgctc tgcgaggctg 17040

gcgggctacc gcgcgggtaa cagatgaggg caagcggatg gctgatgaaa ccaagccaac 17100

caggaagggc agcccacctc tcaaggtgta ptgccttcca gacgaacgaa gagcgattga 17160

ggaaaaaggc gcggcgggcg gcctgagcct gtcggcctac ctgctggcgg tcggccaggg 17220

ctacaaaato acgggcgctc tggactatga gcacgtccgc gagctggccc gcactaatgg 17280

cgacctgggc cgcctggggc gctgctgaa actctggctc accgacgacc cgcgcacggc 17340

gcggttcggt gatgccacga tctcgccct gctggcgaag atcgaagaga agcaggacga 17400
gcttggcaag gtcatgatgg gcgtggtccg cccgagggca gagccatgac ttttttagcc 17460
gctaaaacgg ccgggggggtg cgcgtgattg ccaagcacgt ccccatgcgc tccatcaaga 17520
agagcgactt cgcggagctg gtgaagtaca tcaccgacga gcaaggcaag accgagcgcc 17580
tttgcgacgc tca 17593

<210> 43

<211> 16954

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 43

ccgggctggt tgcctcgcc gctgggctgg cggcgtcta tggcctgca aacggcgaag 60
aaacgcgcgc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgagg 120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gcgcgcgacg tggagctggc 240

cagcctcgca aatcggcgaa aacgcctgat ttacgcgag ttccacag atgatgtgga 300

caagcctggg gataagtgc ctgcggtatt gacacttgag gggcgcgact actgacagat 360

gagggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg ggcacccat 420

tgacatttga gggcgtgtcc acaggcagaa aatccagcat ttgaagggt ttccgccct 480

tttcggcca cgcctaacct gtcttttaac ctgcttttaa accaatattt ataaccttg 540

tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gcgaagggg ggtgcccccc 600

ctctcgaac cctcccgccc cgctaacgcg ggctcccat ccccccagg gctgcgcccc 660

tggcgccga acggcctcac cccaaaaatg gcagcgctgg cagtccttgc cattgcggg 720

atcggggcag taaagggatg ggcgatcagc ccgagcgga cgcgcggaag cattgacgtg 780

ccgcagggtc tggcatcgac attcagcgac cagggtccgg gcagtgaggg oggcggcctg 840

ggtggcgccc tgcccttcac ttggcgctc ggggcattca cggacttcat ggcggggccc 900

gcaattttta cctggggcat tcttggcata gtggtcggg gtgcctgtct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgcottga atatattgac aatactgata 1140

agataatata tcttttatat agaagatcgc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaagcttg aaaccaggga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gccgatgac ttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcagggt tatgccgctc aattcgctgc gtatctcgct tgctgattac gtgcagcttt 1500

ccottcaggc gggattcata cagcggccag ccatocgtca tccatctcac cactgcaag 1560

ggtagacgca ggctcctaag acgccccagc gtcccatag tgcgttcacc gaatacgtgc 1620

gcaacsaccg tcttcgggag actgtcatat gcgtaaaaca gccagcgcgt gcgcgattta 1680

gccccagcat agccccactg ttcgttcatt tccgcgcaga cgatgacgtc actgccccgc 1740

tgtagtcgctg aggttaccca ctggggcctg agttttttta gtgacgtaaa atcgtgttga 1800

ggccaacgcu ctttaalyggy gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccc ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920

taaggcagtg agagcagaga tagcgtgat gtccggcggt gcttttgcg ttaacgacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100

aaatcggtc cgtgcatact atgtttatag ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggtt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220

aattagcttc ttggggtatc ttttaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcacggg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaagggtata taagctggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgcact ttgaacggca 2520

tgatggctgy agcaatctgc tcatgagtga ggccgatggc gtcccttgcct cggaaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640

tcactccatc gacatatcgg attgtcccta taogaatagc ttagacagcc gcttagccga 2700

attggattac·ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccatth aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttccacgc gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgac ttgggagaag cggcaggcg gacaagtgg atgacattgc 2940

cttctgcgtc cggctgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaaataaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacy alyccygcga caagcaggag cgcaccgact 3120

tcttccgcac caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggaagg 3180

ggtcgctggt attcgtgcac ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacg gaccgacttc attgcgata aggtggatta tctggacacc aaggccaccg 3300

gcgggtcaaa tcaggaataa gggacattg ccccgcgctg agtcggggca atcccgcaag 3360

gaggggtaat gaatcggagc tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgteatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatggctc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgccg cgccatcggc cgcggtggag cgttcgcgtc 3600

gtctcgaaac ggaggcgcca ggtttggcga agtcgatgac catcgacacg cgaggaaacta 3660

tgacgaccaa gaagcgaata acccgccggc aggacctggc aaaacaggtc agcaggacca 3720

agcaggccgc gttgctgana caacogaagc agcagatcaa ggaaatgcag ctttcttctgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

cctgttccac cagcgcaac aagaaaatcc cgcgcgaggc gctgcaaac aaggtcatth 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctcggg gccgacgatg 3960
 acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctctatc ggcgagccga 4020
 tcaccttcac gttctacgag ctttgccagg acctgggctg gtogatcaat ggcgggtatt 4080
 acacgaaggc cgagggaatgc ctgtgcgcgc tacaggcgac ggcgatgggc ttcacgtccg 4140
 accgcgttgg gcacctggaa tcgggtgtgc tgctgcaccg ctccgcgtc ctggacogtg 4200
 gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260
 gcyaccacta caggaanatto atatgggaga agtaccgcaa gctgtgcgcg accgcccagc 4320
 ggatgttcga ctatttcagc tcgcaccggg agccgtaccg gctcaagctg gaaaccttc 4380
 gcctcatgtg cggatcggat tcaccccgcg tgaagaagtg gcgcgagcag gtccgcgaag 4440
 cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500
 tgcattgcaa acgctagggc cttgtgggggt cagtccggc tggggggtca cgaccagcg 4560
 ctttactggc atttcaggaa caagcgggca ctgctcgacg caattgcttc gtcagtatc 4620
 gtcggggacg caccggcgcg tctacgaact gccgataaac agaggattaa aattgacaat 4680
 tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcagatt tcccgagat 4740
 ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
 gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccca 4860
 catcgacgcg gagatcattg ggctgtcgggt ctccaacag gaggacggcc ccaaggacgc 4920
 tcacaaggcg catctgtccg gcgttttctg ggagcccgaa cagcgaggcc gaggggtcgc 4980
 cggtatgctg ctgcggggct tgccggcggg ttattgtctc gtgatgatcg tccgacagat 5040
 tccaacggga atctgggtgga tgcgcatctt catcctcgcc gcacttaata ttctgctatt 5100
 ctggagcttg ttgtttattt cggctctacc cctgccgggc ggggtcgccg cgaaggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtctctggggg ctatttgcgg aactgcgggc gtggcgctgt tgggtgtgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga ccgcgaagt gcaacctccc gtgcctctgc tcacctttac 5400
gcctggcga ctggcggcgg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gctgatccgg 5520
aggggttcta acctacttcc ttctgttccg ggggatctcg gcactcgaac ctacagttgt 5580
ttccttactg ggcttttcta gccccagatc tggggtcgat cagccgggga tgcacaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatctgc aacgttccact tctaaagaaa tagccgact cagcttctct agcggcttta 5760
tccagcgatt tcttattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctcgc agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta cctccgcga gatcatccgt 5940
gtlccaacc cggcagotta gttgcrgttc ttcccaatag catcggtaac atgagcaaag 6000
tctgccgctt tacaacggct ctcccgctga cgcgctccc gactgatggg ctgcctgtat 6060
cgagtgggta ttttgtgccc agctgcgggt cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtgtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cggteccgcg tggtttgccc cagcaggcga 6300
aaatcctggt tgatgggtgt tcogaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gtlagtggtt gtcccggtt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaacccgtct atcagggcga tggccacta cgtgaacct 6480

caccaaatc aagttttttt gggctgaggt gccgtaagc actaaatcgg aaccttaag 6540

ggagccccc atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggtttcc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtaacc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gtcattaag agccagttca tggcggttg catgatggcc gtcatgcac tgtactcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgtga aggcgcttcc gaactcgggt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taagsgtgc 7080

aacagcttcc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagtcctt attgaagtcg gaggaaggag 7200

cyygtcaag aggatattct tgcactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcataaggag 7320

tcctttggct ttcaatattc tcaagtatct cgagtttgaa ctatttcctt gtgaaccttt 7380

tattcaccaa tgagcatttg aatgaacatg aatctgagga ctgcaatgc catgaggttt 7440

tcgaataca tccgtagtgc gaaggcttg gccacctgcg ttggttgaat ttagaactg 7500

gcactattga tcacccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtga ctcgagtacc 7740
 attttaatcl atttgtgttt gatcgagacc taataragcc cctacaacga ccatcaaagt 7800
 cgtatagcta ccagttagga agtggaactca aatcgacttc agcaacatct cctggataaa 7860
 ctttaagcct aaactatata gaataagata ggtggagagc ttataccgag ctcccaaatc 7920
 tgtccagatc atggttgacc ggtgcctgga tottctata gaatcatcct tattcgttga 7980
 cctagctgat tctggagtga cccagagggt catgacttga gcctaaaate cgcgcctcc 8040
 accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgcacgg tgactcttcc 8100
 tggcatgcgg agagacggac ggacgcagag agaagggtg agtaataagc cactggccag 8160
 acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggagng cgcgccctc 8220
 cgccccgaag tggaaaggct ggtgtgcccc tcttgacca agaactctatt gcacatcgg 8280
 agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
 gtatagccgt cggcgaaata gcacgcatt aacctaggtg cagaagtcca attgcttccg 8400
 atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc 8460
 gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
 ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggcagc cggcgagac 8580
 cgggaacaca agctggcagt cgaaccatcc ggtgtctctgc actcgacotg ctgaggtccc 8640
 tcagtcctcg gtaggcagct ttgcccgtc tgtecccgcg gtgtgtcggc ggggttgaca 8700
 aggtccttgc gtcagtcaca catttgttgc catattttcc tgctctcccc accagctgct 8760
 cttttctttt ctctttttt tccatcttc agtatattca tcttcccatc caagaacctt 8820
 tatttccctt aagtaagtac tttgtacat ccaactcca tccttcccat ccttattcc 8880
 ttgaaacctt tcagttcgag ctttcccatc tcctgcagc ttgactaaca gctaccgcg 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcca 9000

aaagttcgac agcgtctcug acctgatgca gototcggag ggcgaagaat ctctgtcttt 9060

cagcttcgat gtaggaggcg gtggatatgt cctgcgggta aatagctgcg ccgattggttt 9120

ctacaagat cgttatgttt atcggaactt tgcctcggcc gcgctcccca ttccggaaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtaacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggtcgcgga 9300

ggccatggat gcgatgcgtg cggccgatct tagccagacg agcgggttcg gccattccg 9360

accgcaagga atcggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cyacaccgtc agtgogtoog tegcgaggg 9480

ttctgatgag ctgatgtctt gggccgagga ctgccccgaa gtccggccacc tctgcaacc 9540

ggatttcgcg tcaacaatg tctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct cgcattgggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttggggcgag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcggggctac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca gcaactgtcc 9960

gagggcaaaq aaatagagta gatgcgcacc gggggatcga tccacttaac gttactgaaa 10020

tcatcaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaagaa 10140

gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttccgggttta 10200

cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtcaacgc 10260
cttncaggct cggcgcaaya gaagaatagc ttagcagagc tattttcatt ttccgagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgcctcttggc 10380
tccacgcgac tataatatttg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctggggt cgcaaagata attgcatggt 10500
tcttccttga actctcaagc ctacaggaca cacattcacc gtaggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcacg gttgcctagt gaatgctccg 10620
taacacccaa tacgcgggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttgttt agaggtaatc cllclttcta gctagaagtc ctgctgtact 10740
gtgtaagcgc ccactocaca tctccactcg acctgcaggc atgcaagctt gagattaaaa 10800
tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860
atacatgcct gactttcttt ttccatgat atgatatatg catatgatag atatacaagc 10920
aatcttcttc aaggagtgtg aaattttgtc ctccaggagc aaaaaaagt tttttttat 10980
acatgtttgt acacaagaat agttaccaat ttgcttttgt cttacgtgct gcaagtttat 11040
atogttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100

tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgtat tttactaatt 11160
cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220
ttgattacac tctaaatctc gatgaaccga ccaaaaagag cgaacaaaga aataatcttg 11280
tgcatcgaa tatgatggaa gattttttcc cctttattct aaatgttgac atagcgtgta 11340
tggtatataa acaaaaagaa attgtacaaa ctttcttttt ttctcttttt attttatctc 11400

tatgttctgtg atttggaaatg ccttgatcgt ttctgttacc gtgattggca tggaaagtgt 11460

tgctgcactg gcacacaaa acatcatgca cggctggggt tggggatggc atctttcaca 11520

tcatgaacgc cgtaaagggt cgtttgaagt taacgatctt tatgcctggg tttttgctgc 11580

attatcgatc ctgctgattt atctgggcag tacaggaatg tggccgcctc agtggattgg 11640

cgcaggtatg acggcgtatg gattactcta ttttatgggt cacgacgggc tgggtcatca 11700

acgttggcca ttccgtata ttccacgcaa gggctacctc aaacggttgt atatggcgca 11760

cogtatgcac cagcgcgtca gggcacaaga aggttgtgtt tcttttggct tctctatgc 11820

ggcggccctg tcaaaacttc aggcgcgct cggggaaga catggcgcta gagcgggcgc 11880

tgccagagat gcgcagggcg gggaggatga gccgcaccc gggaagtaag ggccgtacca 11940

gaggcggcca gcagcagcgt taatttttgc ggcgtggctg ttgactgcgc ctgatcccaa 12000

agcttggcgt aatcatggtc atagctgtt cctgtgtgaa atgttatcc gtcacaatt 12060

ccacacaaca tacagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgcgc 12120

taactcaat taattgcgtt gcgctcactg cccgctttcc agtcgggaa cctgtcgtgc 12180

cagctgcatt aatgaatgg ccaagcgcg gggacagcgc gtttgcgtat tgggcaaaag 12240

acaaaaggcg gacattoaac cgattgaggg aggyaaggta aatattgacg gaaattatc 12300

attaaagggt aattatcac gtcaccgact tgagccattt gggaattaga gccagcaaaa 12360

tcaccagtag caccattacc attagcaagg ccggaaacgt caccaatgaa accatcgata 12420

gcagcacgt aatcagtagc gacagaatca agtttgcctt tagcgtcaga ctgtagcgcg 12480

ttttcatogg cattttgggt catagcccc ttattagcgt ttgcactctt ttcataatca 12540

aaatcacggg aaccagagcc accacgggaa ccgcctccct cagagccgcc accctcagaa 12600

cggccacct cagagccacc accctcagag ccgcacacg aaccaccaac acagccggcg 12660

ccagcattga caggaggccc gatctagtaa catagatgac accgcgcgcg ataatttato 12720
ctagtttgcg cgctatatatt tgttttctat cgcgtattaa atgtataatt gogggactct 12780
aatcataaaa acccatctca taaataacgt catgcattac atgttaatta ttacatgctt 12840
aacgtaattc aacagaaatt atatgataat catcgcaaga ccggcaacag gattcaatct 12900
taagaaactt tattgcaaaa tgtttgaacg atcggggatc atccgggtct gtggcgggaa 12960
ctccacgaaa atatccgaac gcagcaagat atcgcggtgc atctcggctt tgcctgggca 13020
gtgcgcgcgc acgcggttga tgtggacgcc gggcccgatc atattgtcgc tcaggatcgt 13080
ggcgttgtgc ttgtcgccg ttgctgtcgt aatgatatcg gcaccttga ccgcctgttc 13140
cgcagagatc ccgtggcgca agaactccag catgagatcc ccgcgctgga ggatcatcca 13200
gccggcgctc cggaaaacga ttccgaagcc caacctttca tagaaggcgg ccgtggaatc 13260
gaaatctcgt gatggcaggt tggcgctcgc ttggtcggtc atttcgaacc ccagagtcgc 13320
gtcagaaga actcgtcaag aaggcgatag aaggcgatgc gctgcgaato gggagcggcg 13380
ataccgtaaa gcacgaggaa gcggtcagcc cattcgcgcg caagctcttc agcaatatca 13440
cyygtagcca ccgctatgto ctgatagcga tccgccacac ccagccggcc acagtcgatg 13500
aatccagaaa agcggccatt ttccaccatg atattcggca agcaggcatc gccatgggtc 13560
acgacgagat catcgcgcgc gggcatgcgc gccttgagcc tggcgaacag ttccggtggc 13620
ggagccctct gatgctcttc gtccagatca tccgtgatga caagaccggc ttccatccga 13680
gtacgtgtc gtccgatgcg atgtttcgtc tgggtgtcga atgggcaggt agccggatca 13740
agcgtatgca gccgccgcct tgcacagcc atgatggata ctttctcggc agggcgaagg 13800
tgagatgaca ggagatcctg ccccgccact tggcccaata gcagccagtc ccttcccgct 13860
tcagtacaaa cgtcgagcac agctggcga ggaacgccc tgggtgccag ccaagatagc 13920

cgcgctgect cgtcctgcag ttcatcagg gcaccggaca ggctgggtctt gacaaaaaga 13980

acggggcgcc cctgcgctga cagccggaac acggcgcat cagagcagcc gattgtctgt 14040

tggtcccagt catagcgaa tagcctctcc acccaagcgg cgggagaacc tgcgtgcaat 14100

ccatcttggt caatcatgcg aaacgatcca gatcgggtgc agattatttg gattgagagt 14160

gaatatgaga ctctaattgg ataccgaggg gaatttatgg aacgtcagtg gagcattttt 14220

gacaagaaat atttgctagc tgatagtgc cttaggcgac ttttgaacgc gcaataatgg 14280

ttttgcagct atgtgcttag ctcatataac tccagaaacc cggcgctgag tggctccttc 14340

aacgttgctgg ttctgtcagt tccaaacgta aaacgggttg tcccgcgta cggcggggg 14400

tcataacgtg actcccttaa ttctcgcctc atgatcagat tgctggtttcc cgccttcagt 14460

ttaaactatc agtggtttgac aggatattt ggcggttaa cetaagagaa aagagcggtt 14520

attagaataa tcggatatatt aaaagggcgt gaaaagggtt atcgttctgt ccatttgtat 14580

gtgcatgcca accacaggtt tcccagatc tggcgccggc cagcgagacg agcaagattg 14640

ggcgcgcccc gaaacgatcc gacagcgcg ccagcacagg tgcgcaggca aattgcacca 14700

acgcatacag cgccagcaga atgccatagt gggcggtgac gtcgttcgag tgaaccagat 14760

cgcgcaggag gcccggcagc accggcataa tcaggccgat gccgacagcg tcgagcgaga 14820

cagtgtcag aattacgatc aggggtatgt tgggtttcac gtctggcctc cggaccagcc 14880

tcgcgtggtc cgattgaacg cgggattct ttatcactga taagttgggt gacatattat 14940

gtttatcagt gataaagtg caagcatgac aaagttgcag ccgaatacag tgatccgtgc 15000

cgccctggac ctgttgaacg aggtcggtgc agacggtctg acgacacgca aactggcgga 15060

acggttgggg gttcagcagc cggcgcttta ctggcacttc aggaacaagc gggcgctgct 15120

cgagcactg gccgaagcca tgctggcgga gaatcatag cattcggtgc cgagagccga 15180

cgacgactgg cgctcatttc tgatcgggaa tgcccgacgc ttcaggcagg cgetgetcgc 15240
ctaccuyat gggcgcgcga tccatgcgcg caccgcaccg gggcaccgc agatggaaac 15300
ggcgcacgcg cagcttgcgt tcctctgcga ggccgggttt tcggccgggg accgcgtcaa 15360
tgcgctgatg acaatcagct acttcactgt tggggccgtg cttgaggagc aggcggcgga 15420
cagcgatgcc ggcgagcgcg gcggcacctg tgaacaggct ccgctctcgc cgcgtgttgcg 15480
ggcgcgcgata gacgccttcg acgaagccgg tccggacgca gcgttcgagc agggactcgc 15540
gggtattgtc gatggattgg cgaaaaggag gctcgttgtc aggaacgttg aaggaccgag 15600
aaagggtagc gattgatcag gaccgctgcc ggagcgcaac ccaactacta cagcagagcc 15660
atgtagacaa catcccctcc cctttccacc cggctcagc gnnctagca ccccgctacg 15720
ggctttttca tgccttcccc tagcgtccaa gcttcacggc cgcgctcggc ctctctggcg 15780
gccttctggc gctcttcgcg ttctcgcgc actgactcgc tgcgctcggc cgttcggctg 15840
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggc tatccacaga atcaggggat 15900
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccagggaaccg taaaaaggcc 15960
gcgttgcgtg cgtttttcca taggctcgcg cccctgacg agcatcaca aaatcgacgc 16020
tcaagtcaga ggtggcgaaa ccgcacagga ctataaagat accaggcggt tccccctgga 16080
agetccctcg tgcgctctcc tgttcgagc ctgcgcctta ccggatacct gtcgctctt 16140
ctcccttcgg gaagcgtggc gcttttcgcg tgcataaccg tgcctcgggg tcattatagc 16200
gattttttcg gtatatccat cctttttcgc acgatataca ggattttgcc aaaggggtcg 16260
tgtagacttt ccttgggtga tccaacggcg tcagccgggc aggataggcg aagtagccc 16320
accgcgcgagc ggggttccct tcttactgt ccctattcg caccctggcg tgcctaacgg 16380
gaalcctgct ctgagaggct ggccggctac cgcggcgta acagatgagg gcaagcggat 16440

ggctgatgaa accaagccaa ccaggaaggg cagcccacct atcaaggtgt actgccttcc 16500
agacgaacga agagcgattg aggaaaaggg ggcgcggccc ggcctgagcc tgcggccta 16560
cctgctggcc gtcggccagg gctacaaaat cacgggctgc gtggactatg agcagctccg 16620
cgagctggcc cgcacaaatg gcgacctggg ccgcctgggc ggccctgctga aactctggct 16680
caccgcagac ccgcgcacgg cgcggttcgg tgatgccacg atcctcgccc tgcggcgaa 16740
gatcgaagag aagcaggacg agcttggcaa ggtcatgatg ggcttggtcc gcccgagggc 16800
agagccatga ctttttttagc cgctaaaacg gccgggggggt gcgcgtgatt gccagcacg 16860
tccccatgcg ctccatcaag aagagcgact tcgcggagct ggtgaagtac atcaccgacg 16920
agcaaggcaa gaccggcgcc ctttgagacg ctca 16954

<210> 44

<211> 16954

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 44

ccyygctggg tgccctcgcc gctgggctgg cgggcgtcta tggccctgca aacgcgccag	60
aaacgcgcgc gaagccgtgt gcgagacacc ggggcgcgcg gcgttgctga tacctgcgcg	120
aaaacttgcc cctcactgac agatgagggg cggacgttga cacttgaggg gcgcactcac	180
ccggcgccgc gttgacagat gaggggcagg ctgcatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgcgag ttcccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gaggggcgcg atccttgaca cttgaggggc agagtgtgga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acagggcagaa aatccagcat ttgcaagggt ttccgcnngt	480
ttttoggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc	600
cttctcgaa cctcccgccc cgctaacgcg ggcctcccat cccccaggg gctgcgcccc	660
tggccgcgca acggcctcac cccaaaaatg gcagcgctgg cagtccttgc cattgccggg	720
atcggggcag taacgggatg ggcgatcagc ccgagcgcca cgcgcggaag cattgacgtg	780
ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcgccctg	840
ggtgcgggcc tgccttcac ttggccgcgc ggggcattca cggacttcac ggcggggcgc	900
gcaattttta ccttgggcat tcttggcata gtggtgcggg gtgcccgtgt cgtgttcggg	960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa	1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag	1080
acgaagcgga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata	1140

agataatata tcttttatat agaagatatc gcggtatgta aggtattcag ggggcaaggc 1200
 ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260
 ctaatgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320
 atgactocaa ottattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380
 agctccaccg attttgagaa cgacagcgac ttccgacca gccgtgccag gtgtgtcttc 1440
 agattcaggt tatgccgtc aattogctgc gtatatcgtc tgtgtattac gtgcagcttt 1500
 ccttcaggc gggattcata cagcggccag ccattccgtc tccatcac cagctcaaag 1560
 ggtgacagca ggctcataag acgcccagc gtgccatag tgcgttcacc gaatacgtgc 1620
 gcaacaaccg tottcggag actgtcatc gcgtaaaaa gccagcgtg gcgcgattta 1680
 gcccgcacat agccccactg ttctgcatt ttccgcaga cgatgacgc actgcccgcc 1740
 tgtatgcgog aggttacga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800
 ggccaacgcc cataatcggg gctgttgccc ggcattcaac gccattuatg gccattatca 1860
 tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgcattgttt 1920
 taaggcagtg agagcagaga tagcgtgat gtccggcggt gcttttgcg ttacgcacca 1980
 ccccgctcgt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
 aaaaacacca tcatacata aatcagtaag ttggcagcat caccataat tgtgtgttca 2100
 aatcggctc cgtcgatact atgttatcgc ccaactttga aaacaacttt gaaaaagctg 2160
 tttcttggtta ttaagggttt tagaatgcaa ggaacagtga attggagtc gtctgtttat 2220
 aattagcttc ttgggtatc tttaaatact gtgaaaaga ggaaggaaat aataatggc 2280
 taaaatgaga atatcaccgg aattgaaaa actgatcgaa aataccgct gcgtaaaaga 2340
 lacggaagga atgtctctg ctaaggtata taagctggtg ggagaaaatg aaaaactata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacotat gatgtggaac gggaaaagga 2460
 catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgcact ttgaacggca 2520
 tgaaggctgg agcaatctgc tcatgagtga ggccgatggc gtcccttgct cggaagagta 2580
 tgaagatgaa caaagccctg aaaagattat cgagctgtat cgggagtcca tcaggctcct 2640
 tcactccatc gacatatacg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
 attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
 cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
 ggaacttgct tttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
 agtaagtggc ttattatgac ttgggagaag cggcagggcg gacaagtgtg atgacattgc 2940
 cttctgcctc cggctgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000
 tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060
 attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
 tcttcgcat caagtgtttt ggctctcagg ccgagggcca cggcaagtat ttgggcaagg 3180
 ggtcgctggt attcgtgtag gccaaagatc ggaataccaa gtacgagaag gacggccaaga 3240
 cggctacagg gaccgacttc attgocgata aggtggatta totggacacc aaggcaccag 3300
 gcggtgcaaa tcaggaataa gggcacattg ccccgcgctg agtcgggca atcccgaag 3360
 gagggtgaat gaatggcagc ttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
 ggttttccgc cgaggatgcc gaaacatcg caagccgcac cgtcatgctg gcgccccgag 3480
 aaaccttcca gtccgtgggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540
 gcgtgcaact ggtccccct gccctgcccg ccgcacatggc cgcctggag cgttccgctc 3600
 gtctcyaca ggaggcgga ggtttggcga agtcgatgac catcgacaag cgaggaaacta 3660

tgacgaccaa gaagcgaaaa accgcggcg aggacctggc aaaacaggtc agcgaggcca 3720
 agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780
 togatattgc gcgcgtggcg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840
 ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggctcattt 3900
 tpcacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccyacyatg 3960
 acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020
 tcacottcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080
 acacgaagcg cgaggaatgc ctgtgcgcc tacaggcgac ggcgatgggc ttacgtccg 4140
 accgcgttgg gcacctggaa tcggtgtcgc tgotgcaccg cttccgcgc ctggacogtg 4200
 gcaagaaaac gtccogttgc caggtcctga tcgacgagga aatcgctgtg ctgtttgctg 4260
 ggcaccacta cagcaaatc atatgggaga agtaaccgaa gctgtgcgc acggcccgac 4320
 ggcgtttcga ctatttcagc tcgcaccggg agcgtatccc gctcaagctg gaaaaccttc 4380
 goctcatgtg cggatcggat tcacccgcg tgaagaagtg gcgcgacgag gtcggcgaa 4440
 cctcggaaga gttgcgagc aacggccttg tggaaacgcg ctgggtcaat gatgacctgg 4500
 tgcattgcaa acgctagggc cttgtggggt cagttccggc tgggggttca gcagccagcg 4560
 ctttactggc atttcaggaa caagcgggca ctgctgcag caottgotto gctcagtac 4620
 gctcgggacg caoggcgcgc tctaogaaat gccgataaac agaggattaa aattgacaat 4680
 tgtgattaag gtcagattc gaogccttg agcggccgac gtgcaggatt tocgcgagat 4740
 ccgattgtcg gccctgaaga aagctocaga gatgttcggg tocgtttacg agcacgagga 4800
 gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgoccta 4860
 catcgacggc gagatcattg ggctgtcggt cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttctg ggagcccgaa cagcgaggcc gaggggtcgc 4980
 cggtagctg ctgcgggggt tgcggggggg ttattgtctc gtgatgatcg tccgacagat 5040
 tccaacggga atctggtgga tgcgcattct catctcggc gcaactaata ttctgcattt 5100
 ctggagcttg ttgtttattt cggctctacg cctgccgggc ggggtcggcg cgacggtagg 5160
 cgctgtgcag ccyctgatgg tctgttcat ctctgccgct ctgctaggta gcccgatacg 5220
 attgatggcg gtctggggg ctatttgcgg aactgccggc gtggcgctgt tgggtgtgac 5280
 accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
 ggcgttcgga accgtgtgca ccgcgaagt gcaacctccc gtgctctgct tcaaccttac 5400
 cgcttgccaa ctggcgggcg yaggaactct gctggttcca gtgctttag tgtttgatcc 5460
 gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggcctg gctgatcgg 5520
 agcgggttta acctactcc ttgtgtccg ggggatctcg cgaactgaac ctacagttgt 5580
 ttcttactg ggctttctca gcccagatc tggggctgat cagccgggga tgcacaggc 5640
 cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
 tgatatctc aacgttcaact tctaaagaaa tagccgact cagcttctc agcggcttta 5760
 tccagcgatt tctattatg tcggcatagt tctcaagatc gacagcctgt caccgttaag 5820
 cgagaaatga ataagaaggc tgataattcg yactctcgc agggagatga tatttgatca 5880
 caggcagcaa cgtctgtca tctttacaat caacatgcta cctccgcga gatcatccgt 5940
 gtttcaaacc cggcagctta gttgcggtc ttccgaatag catcggtaac atgagcgaag 6000
 tctgccgctc tacaacggct ctcccgtga cggcgtccc gactgatggg ctgcctgtat 6060
 cgagtgggta ttttgtccg agctgccggg cggggagctg ttggctggct ggtggcagga 6120
 tatattgtgg tgtaacaaaa ttgacgctta gacaactaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
acugcctggc cctgagagag ttgcagcaag cagtcacacg tggtttgccc cagcaggcga 6300
aaatcctggt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaagggcga aaaaccgctc atcagggcga tggcccacta cgtgaaccat 6480
caccctaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaagggg 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660
cgggcctctt cgtatllacg cccgctggcg aaagggggat gtgctgcaag qcgattaagt 6720
tgggtaacgc caggggtttc ccagtcacga cgttgtaaaa cgaaggccag tgaattcgag 6780
ctcggatacc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840
cgaggagcct cgtcctgtca caactaccaa catggagtac gataaggccc agttccgcca 6900
gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcacg tgtacttcaa 6960
gtacaccaac gctcttctga tccagtcgat catccgtga aggcgtcttc gaattcgggt 7020
aagatccacg tcttcgggaa gccagcgact ggtgaactcc agcgtccctt taaggctggc 7080
aacagctttc tcagccaggc ccagcccaag accgacaagg cctccctcca gaagccgag 7140
aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200
cgggtgcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctgcttgga gcatagggg 7320
tcctttggct ttoaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaatata tccggatgtc gaaggcttgg ggcacotgcg ttggtgaat ttagaacgtg 7500
gcaciatlga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560
tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620
tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680
gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtggc ctcgagtacc 7740
atttaattct atttctgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800
cgtatagcta ccagtggagg agtggactca aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aactatata gaataagata ggtggagagc ttataccgag ctcccaaact 7920
tgtccagatc atgggtgacc yylycctgga tcttctata gaateatctc tattcgttga 7980
cctagctgat tctggagtga cccagagggt catgacttga gctaaaaac cgcgcctcc 8040
accatttcta gaaaaatgtg acgaactcgt gagctctgta cagtgcacgg tgactctttc 8100
tggcatgcgg agagacggac ggcgcagag agaagggtg agtaataagc cactggccag 8160
acagctctgg cggctctgag gtgcagtggc tgattattaa tcggggaccg gcgcgccctc 8220
cgccccgaag tggaaaggct ggtgtgccc tctgtgacca agaactctatt gcctcatcgg 8280
agaatatgga gcttcacga atcaccggca gtaagcgaag gagaatgta agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aaccclayta cagaagtcca attgcttcg 8400
atctggtaaa agattcacga gatagtaoct tctcogaagt aggtagagcg agtaccggc 8460
gcgtaagctc cctaattggc ccatcggca tctgtagggc gtccaaatat cgtgcctctc 8520
ctgctttgcc cgggtgatga aaccggaaag gcgcctcagg agctggccag cgggcgagac 8580
cgggaacaca agctgggaag cgaacctcc ggtgctctgc actcagctg ctgaggtccc 8640

tcagtcacctg gtaggcagct ttgccccgc tgcgcgccg gtgtgtcgcc ggggttgaca 8700
 aggtcggtgc gtcagtcaca catttgtgc catatctcc accagctgct 8760
 cttttctttt cttttctttt tccatcttc agtatattca tcttcccat caagaacctt 8820
 tatttccctt aagtaagta tttgtacat ccatactca tcttcccat ccttattcc 8880
 ttgaacctt tcagttogag ctttccact tcctgcgcg ttgactaaca gctaccgcc 8940
 ttgagcagac atcaccatgc ctgaactcac cgcgcgctt gtcgagaagt ttctgatga 9000
 aagttccac agcgtctccg aactgatga gctctggag gggaagaat ctctgtcttt 9060
 cagcttgat gtaggagggc gtgatatgt cctgcgggta aatagctgcg ccgatgggtt 9120
 ctacaaagt cgttatgttt atcggcactt tgcctcgcc gcgctccga ttcgggaagt 9180
 gcttgacatt ggggaattca gcgagagctt gacctattgc atctccgcc gtgcacagg 9240
 tgtcacgttg caagactgc ctgaaaccga actgcgcgt gttctgcgc cggctgcgga 9300
 ggccatggat gcgctgcgt cggccgatct tagccagag agcgggttcg gccattccg 9360
 accgcaagga atcggtcaat aactacatg gcgtgatttc atatgcgga ttgctgatcc 9420
 ccctgtgtat cactggcaaa ctgtgatga cgcaccgc agtcgctcg tcgcgcaggc 9480
 tctcgatgag ctgatcttt gggccgagga ctgcccga gtcgggcacc tcgtgcaagc 9540
 ggatttcgcg tccaacaatg tctgacgga caatggccg ataacagcg tcattgactg 9600
 gagcgaggcg atgttcgggg attccaata cggagtcgc aacatctct tctggaggcc 9660
 gtggttgct tgtatggag agcagacgc ctacttcgag cggaggaatc cggagcttc 9720
 aggatccgc cggctccgg cgtatctgt ccgatttgt cttgaccaac tctatcagag 9780
 ctgggttgac ggcaatttc atgatgcgc ttgggcgcg ggtgatgcg acgcaatcgt 9840
 ccgatccga gccgggactg tgggggtac acaaatgcg cgcagaagcg cggcgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca gcactcgtcc 9960
gagggcaaaag gaatagagta gatccgacc gggggatcga tocacttaac gttactgaaa 10020
tcacaaaca gcttgacgaa totggatata agatcgttgg tgcgatgtc agctccggag 10080
ttgagacaaa tgytggtcag gatctcgata agatacgttc atttgtcaa gcagcaaaga 10140
gtgccttota gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttcag atacagctca totgcaatgc attaatgcat tgactgcaac ctagttaacg 10260
cttcnaggt ccggcgaaga gaagaatagc ttagcagagc tatttctatt ttccgggagc 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgtctttggc 10380
tcacgcgcac tatatatattg totctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt cgtccaccag cncctgggtt cgcaagata attgcatgtt 10500
tcttccttga actctcaagc ctacaggaca cacattcacc gtaggataaa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcacg gttgcctagt gaatgctccg 10620
taacacccaa tacgcgggc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttggtt aagagtaac cttctttcta gctagaagtc ctctgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt agagataaaa 10800
taaaaagaga agaaaagaaa gtttgtacaa tttctttttg tttatataac atacacgcta 10860
tgtaacatt tagaataagg gggaaaaaat cttccatcat attcgaatgc acaagattat 10920
ttctttgttc gctctttttg gtccggtcat cgagatttag agtgaatca aagatactgt 10980
catctcgaga ggtgtgcaca ggctgctgtt tgocaaattg gatgtttgcc gaattagtaa 11040
aatacgcaag caktctctac ctttcgcctc ctttttcta attctcccaa agactaaatg 11100
aggaaagata aaggacaaag aaatgtaaa gacaaagaa ttgaaaacga tataaacttg 11160

cagcacgtaa gaccaaaagca aattggtaac tattcttgtg tacaaacatg tataaaaaaa 11220

aacttttttt tgcctctgga ggacaaaatt tcaaactcct tgaagaagat tgcttgtata 11280

tctatcatat gcatatatca tatogatgga aaaagaaagt caggcatgta ttataaaaaa 11340

gaagaatgtg ccatgcttcc gaatttcttt tcactttott ttctttatct attttaatct 11400

catgttgttg atttggaatg ccctgacgtt ttctgttacc gtgattggca tggaaagtgt 11460

tgctgcactg gcacacaaat acatcatgca cggctggggt tggggatggc atctttcaca 11520

tcatgaacog cgtaaaggtg cgtttgaagt taacgatctt tatgcctgg tttttgtgc 11580

attatogac ctgctgattt atctgggocg tacaggaatg tggcgcctcc agtgattgg 11640

cgcaggtagt acggcgtagt gattactcta ttttatgggt cagcaggggc tgggtgcatca 11700

acgttggcca ttccgtata ttocacgaa gggctacctc aaacggttgt atatggcga 11760

ccgtatgcat caogcogtca ggggcaaaga aggttgtgtt tcttttggtt tctctatgc 11820

gccgcccctg tcaaaacttc aggcgaogct cgggaaaga catggcgcta gacgggggc 11880

tgccagagat gcgcaggcgg gggaggatga gcccgcatcc gggaagtaag ggcotgacca 11940

gaggcgcca gcagcagcgt taatttttcc ggogtggctg ttgactgcgc ctgacccaa 12000

agcttggcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc gctcacaatt 12060

ccacacaaca tacgagocgg aagcataaag tgtaaagcct ggggtgccta atgagtgcgc 12120

taactcacat taattgcgtt ggcctcactg ccgcctttcc agtcgggaaa cctgtcgtgc 12180

cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggccaag 12240

acaaaagggc gacattcaac cgattgaggg agggaaggta aatattgacg gaaattatc 12300

attaaagggt aattatcacc gtcacogact tgagccattt gggaattaga gccagcaaaa 12360

tcaccagtag caccattacc attagcaagg ccggaaacgt caccaatgaa accatogata 12420

gcagcaccgt aatcagtagc gacagaatca agtttgcctt tagcgtcaga ctgtagcgcg 12480
ttttcaccgg cattttcgggt catagcncnc ttattagcgt ttgccatctt ttcataatca 12540
aatcaccgg aaccagagcc accaccggaa ccgcctccct cagagccgcc accctcagaa 12600
ccgccacct cagagccacc acctccagag ccgccaccag aaccaccacc agagccgcgcg 12660
ccagcattga caggaggccc gatctagtaa catagatgac accgcgcgcg ataattatc 12720
ctagtttgcg cgtatatatt tgttttctat cgcgtattaa atgtataatt ggggactct 12780
aatcataaaa acccatctca taataacgt catgcattac atgttaatta ttacatgctt 12840
aacgtaatto aacagaaatt atatgataat catcgcaaga ccggcaacag gattcaatct 12900
taagaaaactt tattgccaaa tgtttgaaag atcgggggacn atccggctct gtggcgggaa 12960
ctccacgaaa atatcgaac gcagcaagat atcgcggtgc atctcggctc tgcttgggca 13020
gtcgcgcgcg acgcggttga tgtggacgcc gggcccgatc atattgtcgc tcaggatcgt 13080
ggcgttgtgc ttgtcggcgc ttgctgtcgt aatgatatcg gcaccttcga ccgcctgttc 13140
cgcagagatc ccgtgggcga agaactccag catgagatcc ccgcgttgga ggatcatcca 13200
gccggcgtcc cggaaaacga ttccgaagcc caacctttca tagaaggcgg ccgttgaatc 13260
gaaatctcgt gatggcaggt tgggcgtcgc ttgttggttc atttcgaacc ccagagtccc 13320
gctcagaaga actcgtcaag aaggcgaag aaggcgatgo gctgcgaatn gggagcgcgcg 13380
ataccgtaaa gcacaggaa gcggtcagcc cattcgcgcg caagctcttc agcaatatca 13440
cgggttagcca acgctatgtc ctgatagcgg tcgcgcacac ccagccgcc acagtogatg 13500
aatccagaaa agcggccatt ttccaccatg atattcggca agcaggcatc gccatgggtc 13560
acgacgagat catgccgctc gggcatgccc gccttgagcc tggcgaacag ttccgctggc 13620
gcgagccctt gatgctcttc gtccagatca tctgatcga caagacgcgc ttccatcca 13680

gtacgtgctc gctcgtatcg atgtttcgtc tggtagtcga atgggcaggt agccggatca 13740
agcglatgca gcgcgcgcgt tgcctacgcc atgtaggata cttttcggc agggacaagg 13800
tgagatgaca ggagatccctg ccccggaact tcgccaata gcagccagtc ccttcccgct 13860
tcagtgcaca cgtcgagcac agctgcgcaa ggaacgcccg tcgtggccag ccacgatagc 13920
cgcgctgcct cgtctcgagc ttoattcagg gcaccggaca ggtcggtctt gacaaaaaga 13980
accggggccc cctgcgtgca cagccggaac accggggcat cagagcagcc gattgtctgt 14040
tgtgccagc catagccgaa tagcctctcc acccaagcgg ccggagaacc tgcgtgcaat 14100
ccatcttgtt caatcatcg aaacgatcca gatccggtgc agattatttg gattgagagt 14160
gaatatgaga ctctaatgg ataccgagg gaatttatgc aacgtcagtg gagcattttt 14220
gacaagaaat attgtctagc ttagatgac cttaggcgac ttttgaaagc gcaataatgg 14280
ttttgcagct atgtgcttag ctcatataac tcagaaaacc cgcggctgag tggctccttc 14340
aacgttgccg ttctgtcagt tccaaacgta aaacggcttg tcccgcgtca tcggcggggg 14400
tcataacgtg actcccttaa ttctccgtc atgatcagat tgtgtttcc cgccttcagt 14460
ttaaactatc agtgtttgac aggatattt ggccggtaaa cctaagagaa aagagcggtt 14520
attagaataa tcggatattt aaaagggcgt gaaagggtt atcgttgtt ccatttgtat 14580
gtgcattgcca accacagggt tcccagatc tgggcgggc nacgcagacg agcaagattg 14640
ggcgcggccc gaaacgatcc gacagcgccg ccagcacagc tgcgcaggca aattgcacca 14700
acgcatacag cgcacgcaga atgccatagt gggcggtgac gtcgttcgag tgaaccagat 14760
cgcgaggagg gcccggaagc accggcataa tcaggccgat gccacagcg tcgagcgcca 14820
cagtgctcag aattacgac aggggtatgt tgggtttcac gtcggcctc cggaccagcc 14880
tccgctggtc cgattgaacg cgcggattct ttatcactga taagtgttg gacatattat 14940

gtttatcagt gataaagtgt caagcatgac aaagttgcag ccgaatacag tgatccgtgc 15000
cgccctggac ctggtgaaag aggtcggcgt aagcggctctg acgacacgca aactggcgga 15060
acggttgggg gttcagcagc cggcgcttta ctggcacttc aggaacaagc gggcgctgct 15120
cgacgcactg gccgaagcca tgctggcgga gaatcatacg cattcggtgc cgagagccga 15180
cgacgactgg cgctcatttc tgatcgggaa tgcccgagc ttcaggcagg cgctgctcgc 15240
ctaccggat ggcgcgcga tccatgccg cagcgagccg ggcgcaccgc agatggaaac 15300
ggccgacgcg cagcttcgct tctctcgca ggcgggtttt tcggccgggg acgccgtcaa 15360
tgcgctgatg acaatcagct acttcaactgt tggggccgtg cttgaggagc aggccggcga 15420
cagcgatgcc ggcgagcgcg ggcgcaccgt tgaacaggct ccgctctcgc cgctggtgcg 15480
ggccgcgata gacgccttcg acgaagccgg tccggacgca gcgttcgagc agggactcgc 15540
ggtgattgtc gatggattgg cgaaaaggag gctcgtgtgc aggaacgttg aaggaccgag 15600
aaaggggtgac gattgatcag gaccgctgcc ggagcgcaac ccactcata cagcagagcc 15660
atgtagacaa catccctcc cctttccac cgctcagac gcccgtagca gcccgctacg 15720
ggctttttca tgccctgcc tagcgtccaa gctcacggc cgcgctcggc ctctctggcg 15780
gcctctggc gctcttcgc tctctgcct actgactcgc tgcgctcggc cgttcggctg 15840
cggcgagcgg tatcagctca ctcaagcgc gtaatacggc tatccacaga atcaggggat 15900
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaacg taaaagggcc 15960
gcgttgctgg cgtttttcca taggctcgc cccctgacg agcatcaaaa aaatcgagc 16020
tcaagtcaga ggtggcgaaa ccgacagga ctataaagat accaggcggt tcccctgga 16080
agctccctcg tgcgctctcc tgttcagacc ctgcgctta ccgataact gtccgccttt 16140

ctcccttcgg gaagcgtggc gcttttcgcg tgcataaccc tgcttcgggg tcattatagc 16200
gatttttttcg gtatatccat cctttttcgc acgatataca ggattttgcc aaagggttcg 16260
tgtagacttt ccttggtgta tccaacggcg tcagccgggc aggataggtg aagtaggccc 16320
accgcgagc ggggtgttct tcttcaactgt cccttattcg cacctggcgg tgctcaacgg 16380
gaatcctgct ctgcgaggct ggccgggctac cgcggcgcta acagatgagg gcaagcggat 16440
ggctgatgaa accaagccaa ccagggaagg cagccacact atcaaggtgt actgccttcc 16500
agacgaanga agagccattg aggaaaaggc ggccggcgcc ggcatgagcc tgcgggcta 16560
cctgctggcc gtcggccagg gctacaaaat caccggcgtc gtggactatg agcagctccg 16620
cgagctggcc cgcacatag cgcacctggg ccgcctgggc ggctgctga aactctggct 16680
cacgcagac ccgcgcacgg cgcggttcgg tgatgccacg atcctcgccc tgctggcgaa 16740
gatcgaagag aagcaggacg agcttgcaa ggtcatgatg ggctggtcc gcccgaggcc 16800
agagccatga ctttttttagc cgctaaaacg gccggggggt gcgcgtgatt gccaaagucg 16860
tccccatgog ctccatcaag aagagcgact tcgcggagct ggtgaagtac atcacgcagc 16920
agcaaggcaa gaccgagcgc ctttcgagc ctca 16954

<210> 45

<211> 19491

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (18970)..(18970)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (19178)..(19178)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (19269)..(19269)

<223> n is a, c, g, or t

<400> 45

agcttgggtac cgagctcgga tccactagta acggccgcca gtgtgttgga attgcacctt	60
gacggccaggt gaattcgagc tcgggtaccgg gggatctttc gacctgaaa tacgtcgagc	120
ctgtccgctt tggaagcggc gaggagcttc gtctgtgac aactaccac atggaactag	180
ataagggcca gttccgccag ctcatataaga gccagttcat gggcggttggc atgatggccg	240
tcattgcattt gtacttcaag tacaccaacg ctctctgat ccagtcgac atccgctgaa	300
ggcgcttttg aatctgggta agatccacgt ctctgggaag ccagcgactg gtgacctcca	360
gcgtcccttt aaggctgcca acagctttct cagccagggc cagcccaaga ccgacaaggc	420
ctccctccag aacgcccaga agaactggag ggggtgtgtc aaggaggagt aagctcctta	480
ttgaagtggc aggacggagc ggtgtcaaga ggatattctt cgactctgta ttatagataa	540
gatgatgagg aattggaggt agcatagctt catttggatt tgetttccag gctgagactc	600
tagcttggag catagagggc cctttggctt tcaatattct caagtattct gagtttgaac	660
ttattccctg tgaccctttt attcaccaat gaggattgga atgaacatga atctgaggac	720
tgcaatogcc atgaggtttt cgaatacat ccggtatgtc aaggcttggg gcaactgcgt	780
tgggtgaatt tagaacgtgg cactattgat catccgatag ctctgcaaag ggcgttgca	840
aatgcaagtc aaacgttgtc agcagttcca ggtggaatgt tatgatgagc attgtattaa	900

atcaggagat atagcatgat ctctagttag ctcaccacaa aagtcagacg gcgtaaccaa 960
 aagtcacaca acacaagctg taaggatttc ggcacggcta cggaagacgg agaagccacc 1020
 ttcagtggac tcgagtacca ttttaattcta ttgtgtttg atcgagacct aatacagccc 1080
 ctacaacgac catcaaagtc gtatagctac cagtgaggaa gtggactcaa atcgacttca 1140
 gcaacatetc ctggataaac ttttaagccta aactatacag aataagatag gtggagagct 1200
 tataccgagc tcccaaatct gtccagatca tgggtgaccg gtgcctggat ctctctatag 1260
 aatcatcctt attcgttgac ctagctgatt ctggagtgc cagagggtc atgacttgag 1320
 cctaaaatcc gccgcctcca ccatttctag aaaaatgtga cgaactcgtg agctctgtac 1380
 agtgaccggt gactcttttc ygcctggga gagaaggacg gacgcagaga gaagggctca 1440
 gtaataagcc actggccaga cagctctggc ggctctgagg tgcagtggat gattattaat 1500
 cggggaccgg ccgcccctcc gccccgaagt ggaaggctg gtgtgccctt cgttgaccaa 1560
 gaatctattg catcatcgga gaatatggag ctccatcgaa tcaccggcag taagcgaagg 1620
 agaattgtaa gccaggggtg tatagccgtc ggcgaaatag catgccatta acctaggtag 1680
 agaagtccaa ttgcttccga tctggtaaaa gattcacgag atagtaacct ctccgaagta 1740
 ggtagagcga gtaccggcg cgtaagctcc ctaattggcc catccggcat ctgtagggcg 1800
 tccaaatc gtgcctctcc tgccttggcc ylytatgaa accggaaagg ccgctcagga 1860
 gctggccagc ggccgagacc gggaacacaa gctggcagtc gaccoatccg gtgtctgca 1920
 ctgcacctg tgaggctcc cagtcctgg taggcagctt tgccccgtct gtccgccgg 1980
 tgtgtcggcg ggggtgacaa ggtcgttgcg tcagtccaac atttgttgc atattttcct 2040
 gcttccca ccagctgctc ttttttttc tcttttttt ccatcttca gtatattcat 2100

ctcccatcc aagaaccttt attccccccta agtaagtact ttgctacatc catakccat 2160
 cctcccatc ccttatctct ttgaaccttt cagttcgagc ttcccaactt catcgagct 2220
 tgactaacag ctaccccgtc tgagcagaca tcaccatgct gtcgagctg cagtcaatca 2280
 gcgtcaaggc ccgcgcgctt gaactagccc gcgacatcac gcggcccaaa gtctgcctgc 2340
 atgctcagcg gtgctcgta gttcggctgc gaggggcagc accacagaca gagggggcgc 2400
 tgggaacctg gcaggctgcc ggcgcgggcg atgagcacag cgcgatgta gcactccagc 2460
 agcttgacgc ggcctatgca gagcgtcgtc ccggcgcaa acgggagcag ctgtcatacc 2520
 aggtgcgcgc cattgcagca tcaattggcg tgcaggcat tgccatcttc gccacctacc 2580
 tgagatttgc catgcacatg accgtgggcg gcgcagtgcc atgggtgaa gtggctggca 2640
 ctctcctctt ggtggttggg ggcgcgctcg gcctggagat gtatgccgcg tatgcacaca 2700
 aagccatctg gcctagctcg cctctgggct ggctgctgca caagagccac cacacacctc 2760
 gcactggacc ctttgaagcc aacgacttgt ttgcaatcat caatgyactc coegccatgc 2820
 tctgtgtac ctttggcttc tggctgccc acgtcctggg ggcgccctgc tttggagcgg 2880
 ggctgggcat caagctatcc ggcctggcat atatgtttgt acagctgggc ctggtgcaca 2940
 ggctgttcc caccggggcc atcgctggcc tgccctacat gaagcgctgc acagtggccc 3000
 accagctaca ccacagcggc aagtacggcg gcgcgccctg gggatatgtc ttgggtccac 3060
 aggagctgca gcacattcca ggtgcggcgg aggaggtgga gcgactggtc ctggaactgg 3120
 actggtccaa gcggtagggt gcggaaccag gcacgtgggt ttcaacctc atgcctgtga 3180
 taagtgctgg cttagcgat gcgtgtgaga cgggtatgtc acggtcgact ggtctgatgg 3240
 ccaatggcat cggccatgto tggctcatcc gggctgggtg cctgggtgaa ggtgatgcac 3300
 atcatcatgt gcgggtggag gggctgccc agtgtgggct gaactggagc agttgtccag 3360

gctggcggttg aatcagtgag ggtttgtgat tggcggttgt gaagcaatga ctccgcccat 3420
attctatttg tgggagctga gatgatggca tgcttgggat gtgcatggat catggtagtg 3480
cagcaaaacta tattcaccta gggctgttgg taggatcagg tgaggccttg cacattgcat 3540
gatgtactcg tcatggtgtg ttggtgagag gatggatgtg gatggatgtg tattctcaga 3600
cgtagacett gactggaggc ttgatcgaga gactgggccc taltctttga gggggaggc 3660
tcgtgccaga aatggtgagt ggatgactgt gacgctgtac attgcaggca ggtgagatgc 3720
actgtctega ttgtaaaata cattcagatg caagcttggc gtaatcatgg tcatagctgt 3780
ttcctgtgtg aaattgttat ccgctcaca ttccacacaa catcagagcc ggaagcataa 3840
agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaatgtcg ttgcgctcac 3900
tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcy 3960
cggggagagg cggtttgcgt attgggccaa agacaaaagg gcgacattca accgattgag 4020
ggagggaagg taaatatgga cggaatttat tcattaaagg tgaattatca ccgtcacgca 4080
cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta ccattagcaa 4140
ggccggnaac gtaccaatg aaacctoga tagcagcacc gtaatcagta gcgacagaa 4200
caagtttgcc tttagcgtca gactgtagcy cgttttcac gccattttcg gtcatagcc 4260
ccttattagc gtttgcacat ttttcataat caaaatcac ggaaccagag ccaccacgg 4320
aacgcctcc ctacagagcg ccaccctcag aaccgccacc ctacagacca ccaccctcag 4380
agcgccacc agaaccacca ccagagccgc gcgcagcatt gacaggagcg ccgatctagt 4440
aacatagatg acaccgcgcg cgataattta tctagtttg cgcgctatat ttgtttttc 4500
atcgcttatt aaatgtataa ttgogggact ctaatcataa aaacccatct cataaataac 4560
gcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa ttatatgata 4620

atcatcgcaa gacgggcaac aggattcaat cttagaacac ttatttgcca aatgtttgaa 4680
 cgcacgggga tcatcgggt ctgtggcggg aactccacga aaatatccga acgcagcaag 4740
 atatcgcggt gcattctcggg ctgcctggg cagtcgccc gacgcgcgtt gatgtggaag 4800
 cggggcccca tcatattgtc gtcaggatc gtggcgttg gcttgtcggc cgttgctgtc 4860
 gtaatgatat cggcaccttc gacgcctgt tccgcagaga tccgtggg gaagaactc 4920
 agcatgagat cccgcgcgtg gaggatcac cagccggcgt cccggaaaac gattccgaag 4980
 cccaaacctt catagaagc ggcgggtggaa tcgaatctc gtgatggcag gttgggcgtc 5040
 gcttggtcgg tcatttgaa cccagatgc ccgctcagaa gaactcgtc agaaggcgat 5100
 agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg aagcggtcag 5160
 cccattcgcc gccaaagtct tcagcaatat caccggtagc caacgctatg tcctgatagc 5220
 ggtccgccac acccagcggc ccacagtga tgaatccaga aaagcggcca tttccacca 5280
 tgatattcgg caagcaggca tcgcatggg tcacgacgag atcatcgccg tcggscatg 5340
 gcgccttgag cctggcgaac agttcggctg gcgcgagccc ctgatgctct tcgtccagat 5400
 catcctgate gacaagacg gcttccatcc gactacgtc tcgctcgatg cgatgtttcg 5460
 cttggtggtc gaatgggcag gtagccgat caacgctatg cagccgcgc attgcatcag 5520
 ccctgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc tgccccggca 5580
 cttoGCCaa tagcagccag tcccttccc ctccagtac aacgtcgagc acagctgcgc 5640
 aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctgctcctg agttcattca 5700
 gggcacggga caggtcgtc ttgacaaaa gaaccggggc cccctgcgct gacagccgga 5760
 acacggcggc atcagagcag ccgattgtct gttgtgccc gtcatacgcc aatagcctct 5820
 ccaccaagc ggcgggagaa cctgcgtgca atccatcttg ttcaatcatg cgaacgac 5880

cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt ggataccgag 5940
gggaattlat ggaacgtcag tggagcattt ttgacaagaa atatttqcta gctgatagtg 6000
accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt agctcattaa 6060
actccagaaa ccgcggctg agtggtctct tcaacgttgc ggttctgtca gttccaaacg 6120
taaaacggct tgtcccgct catcgccggg ggtoataacg tgactccctt aattctcgc 6180
tcatgatcag attgtcgctt ccgccttca gtttaaacta tcaagtgttg acaggatata 6240
ttggcgggta aacctagag aaaagagcgt ttattagaat aatcgatat ttaaaagggc 6300
gtgaaaaggt ttatcgttc gtccatttgt atgtgcctgc caaccacagg gttccccaga 6360
tctggcgccg gccagcgaga cgaycaagat tggccgcgcg oogaacgat nngacagcgc 6420
gccagcaca ggtgcgcagg caaattgcac caacgcatac agcgccagca gaatgcata 6480
gtggggcgtg acgtcgttc agtgaacag atcgccagg aggcgccgca gcaccggcat 6540
aatcaggcgc atcgccagac cgtcgagcgc gacagtgcgc agaattcga tcaggggtat 6600
gttggttttc acgtctggcc tcgggaccag cctccgctgg tccgattgaa cgcgcggtat 6660
ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt gtcagcatg 6720
acaaagtgc agccgaatac agtgatcgt gcgcgccctg acctgttgaa cgaggtcggc 6780
gtagacggtc tgacgacacg caaactggcg gaacyyllgg gggttcagca gcggcgctt 6840
tactggcact tcaggaacaa gcgggcgctg ctgcagcgc tggccgaagc catgctggcg 6900
gaqaatcata cgcattcggt gccgagagcc gacgacgact ggcgctcatt tctgatcggg 6960
aatgccgca gcttcaggca ggcgctgctc gctaccgcg atggcgcgcg catccatgcc 7020
ggcacgcgac cgggcgcacc gaagatggaa acggccgacg cgcagcttcg ctctctctgc 7080
gaggcgggtt tttcggcgcg ggacgcgctc aatgcgctga tgacaatcag ctacttcaat 7140

gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg cggcggcacc 7200
gttgaacagg ctccgctctc gcgctgttg cgggcccga tagacgcctt cgacgaagcc 7260
ggtcgggacg cagcgttcga gcagggactc gcggtgattg tcgatggatt ggcgaaaagg 7320
aggctcgttg tcaggaaagt tgaaggaccg agaaaagggtg acgattgatc aggaccgctg 7380
ccggagcgca acccactcac tacagcagag ccatttagac aacatccctt ccccccttc 7440
accggtcag acgcccgtag cagccccta cgggcttttt catgcoctgc cctagcgtcc 7500
aagctccag ccgcgctcg gcctctctgg cggccttctg gcgtctcttc gcttcctgc 7560
tcaactgactc gctgcgctcg gtcttctggc tgcggcgagc ggtatcagct cactcaagg 7620
cggtaatacg gttatccaca gaalccgggg ataagcagg aagaacatg tqagcaaaag 7680
gccagcaaaa ggcaggaaac cgtaaaaagg ccgcttctgt ggcgttttct cataggctcc 7740
gccccctga ccagcatcac aaaaatcgac gctcaagtca gaggtagcga aaccgcacag 7800
gactataaag ataaccaggc tttccctcg gaagctccct cgtgcgctct cctgttccga 7860
ccctgcgct tacaggatac ctgtccgctt ttctcccttc gggagcgtg gcgcttttc 7920
gctgcataac cctgcttcgg ggtoattata gcgatttttt cggatatcc atccttttc 7980
gcacgataa caggattttg ccaaagggtt cgtgtgact ttcttgggt tatccaagg 8040
cgtcagccgg gcaggatagg tgaagtaggc caaccgcga ggggtgttc ctcttact 8100
gtcccttatt cgcacctggc ggtgtcaac gggaactctg ctctgcagg ctggccgct 8160
accgccggg taacagatga gggcaaggc atggctgatg aaaccaagcc aaccagggaag 8220
ggcagccac ctatcaagggt gtactgcctt ccagcgaac gaagagcgat tgaggaaaag 8280
ggcgccggg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca gggctacaa 8340
atcacggcg tcgtggacta tgagcacgtc ccgagctgg ccgcaccaa tggcgacctg 8400

ggcgcgcctgg ggggcctgct gaaactctgg ctcaccgacg acccgcgac ggcgcggttc 8460

ggatgatgca cgcctcctcgc cctgctggcg aagatogaag agaagcagga cgcgcttgcc 8520

aaggctcatga tggcgctggt ccgcccagg gacagccat gactttttta gccgctaata 8580

cggccggggg gtgcgcgtga ttgcaagca cgtcccatg cgctccatca agaagagcga 8640

cttcgcggag ctggtgaagt acatcaccga cgagcaaggc aagaccgagc gcccttgaga 8700

cgctcaccgg gctggttgcc ctgcgcgctg ggctggcgcc cgtctatggc cctgcaaacg 8760

cgccagaaac gccctgaag ccgtgtgcga gacaccgagg ccgcccggct tgtggatacc 8820

tgcgggaaaa cttggccctc actgacagat gggggcgga cgttgacact tgaggggccg 8880

actcaccggg cgcggcggtt acagatgagg ggcayyctgc atttcggcgg gcgaagtgga 8940

gctggccago ctgcataac ggcgaaaacg cctgatttta cgcgagtttc ccacagatga 9000

tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc gcgactactg 9060

acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgcctgacga tgaggggcgc 9120

acctattgac atttgagggg ctgtccacag gcagaaaac cagcatttgc aagggtttcc 9180

gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca atatttataa 9240

acctgtttt taaccagggc tgccgcctgt ggcgctgacc ggcacgcgc aaggggggtg 9300

cccccccttc tgaacccctc ccggcccgcct aacgcgggac tcccatcccc ccaggggctg 9360

cgccctcctg ccgcgaacgg cctcacccca aaaatggcag cgctggcagt ccttgccatt 9420

gccgcgctgc ggcagtaac gggatggcg atcagccga ggcgcagcc ccgaagcatt 9480

gacgtgcgcg aggtgctgac atcgacattc agcgaccagg tgccgggcag tgaggggcgc 9540

ggcctgggtg ggcgcctgac cttaacttcg gcgctgggg cattcagga cttcatgacg 9600

gggocggcaa tttttacctt gggcattctt ggcatagtgg tcgocgggtgc cgtgctcgtg 9660
ttcgggggtg cgataaacc agogaacat ttgaggtgat aggtaagatt ataccgaggt 9720
atgaaaacga gaattggacc ttacagaat tactctatga agcgccatat ttaaaaagct 9780
accaagacga agaggatgaa gaggatgagg aggcagattg cttgaatat attgacaata 9840
ctgataagat aatatatctt ttatatagaa gatatgccg tatgtaagga tttcagyyyy 9900
caaggcatag gcagcgcgt tatcaatata tctatagaat gggcaagca taaaaacttg 9960
catggactaa tgottgaan ccagacaaat aaccttatag cttgtaaat ctatcataat 10020
tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgcc gatgactttg 10080
tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg tgccaggtgc 10140
tgccctcagat tcaggttatg ccgctcaatt cgtgcgtat atcgttctgt gattacgtgc 10200
agctttccct tcaggcggga ttcatacagc ggccagccat ccgtcatcca tatcaccagc 10260
tcaaaggggtg acagcaggct cataagacgc ccagcgtcg ccatagtgog ttcaccgaat 10320
acgtgcgcaa caaccgtctt ccggagactg toatacgcgt aaaacagcca gcgctggcgc 10380
gatttaguuc cgacatagcc ccactgttgc tccatttccg cgcagacgat gacgtcactg 10440
cccgctgta tgcgcgaggt taccgactgc ggcctgagtt ttttaagtga cgtaaaatcg 10500
tgttgaggcc aacgcccata atcggggctg ttgccgggca tccaacgcca ttcattggcca 10560
tatcaatgat tttctggtgc gtaccgggtt gagaagcgtt gtaagtgaac tgcagttgcc 10620
atgttttacg gcagtgaag cagagatagc gctgatgtcc ggcgtgtctt ttgcgttac 10680
gcaccacccc gtcagttagct gaacaggagg gacagctgat agacacagaa gccactggag 10740
cacctcaaaa acaccatcat acactaaatc agtaagttgg cagcatcacc cataattgtg 10800
gtttcaaat cgggtccylc gatactatgt tatacgcaa ctttgaanaa aactttgaaa 10860

aagctgtttt ctggtattta aggtttttaga atgcaaggaa cagtgaattg gagttcgtct 10920
tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa ggaataata 10980
aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat accgctgcgt 11040
aaaagatacg gaaggaaatgt ctctgctaa ggtatataag ctggtgggag aaaatgaaaa 11100
cctatattta aaaatgacgg acagccggta taaagggacc acctatgatg tggacggga 11160
aaaggacatg atgctatggc tggaggaaa gctgcctgtt ccaaaggctc tgcactttga 11220
acggcatgat gctgggaga atctgctcat gactgaggcc gatggcgtcc tttgctcgga 11280
agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatcggg agtgcacag 11340
gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag acagccgctt 11400
agccgaattg gattacttac tgaataacga tctggccgat gtggattgag aaaactggga 11460
agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaga cggaaaagcc 11520
cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct tttgaaaga 11580
tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcgaca agtggtatga 11640
cattgccttc tgcgtccgtt ccatcaggga ggaatccggg qaagaacagt atgtcgagct 11700
attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaaatatt atattttact 11760
ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag caggagcgca 11820
ccgactctt cgcacatcaag tgttttggct ctacggcga ggccacggc aagtatttgg 11880
gcaagggttc gctggtatto gtgcagggca agattcggaa taccaagtac gagaaggagc 11940
gccagagggt ctacgggacc gacttcattg ccgataaggt ggattatctg gacaccaagg 12000
caccaggcgg gtcaaatcag gaataagggc acattgccc ggcgtgagtc ggggcaatcc 12060
cgcaaggagg gtgaatgaat cggacytttg acgggaagga atacaggcaa gaactgatcg 12120

acgcgggggtt ttccgccgag gatgccgaaa ccatacgcaag ccgcaccgtc atgcgtgcgc 12180
cccgcgaaac ctccagtcgc gtccggctoga tgggccagca agctacggcc aagatcgagc 12240
gcgacagcgt gcaactggct cccoctgccc tgcccgcgcc atcgcccgcc gtggagcgtt 12300
cgcgtcgtct cgaacaggag gcggcaggtt tggcgaagtc gatgaccatc gacacgcgag 12360
gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa caggtcagcg 12420
aggccaagca ggcgcgcttg ctgaaacaca cgaagcagca gatcaaggaa atgcagcttt 12480
ccttgttcga tattggcgcc tggccgcaca ccatcgagc gatgccaaac gacacggccc 12540
gctctgccct gtaccacagc cgcaacaaga aaatcccgcc cgagcgctgc caaacaagg 12600
tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgtcgag ctgcgggcgc 12660
acgatgacga actggtgttg cagcaggtgt tggagtacgc gaagcgcacc cctatcggcg 12720
agccgatcac ctccacgttc tacgagcttt gccaggacct gggctggtcg atcaatggcc 12780
ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg atgggcttca 12840
cgtccgacgc cgttgggcac ctggaatcgg tctcgtctgt gcaccgcttc cgcgtcctgg 12900
accgtggcaa yaaaaagtc cgttgccagg tcttgatcga ccaggaaatc gtctgtctgt 12960
ttgctggcga ccactacagc aaatccatct gggagaagta ccgcaagctg tcgccgacgg 13020
cccgacggat gttcgactat ttcagctcgc accgggagcc gtaccgcctc aagctggaaa 13080
ccttccgctc catgtcgga tcggattcca cccggtgaa gaagtggcgc gacgaggtcg 13140
gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg gtcaatgatg 13200
acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggtggg ggttcagcag 13260
ccagcgcctt actggcattt caggaaacag cgggcactgc tcgacgcact tgcctcgcgc 13320
agtatcgcgc gggagcgcgc gcgcctccta cgaactgcgc ataacagag cattaataatt 13380

gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc aggatttccg 13440
cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttccgggtccg tttacgagca 13500
cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgcgc tggcattcgg 13560
cgctacatc gacggcgaga tcattgggct gtccgtcttc aaacaggagg accgccccaa 13620
ggacgctcac aaggcgcatc tgtccggcgt ttctgtggag cccgaacagc gaggccgagg 13680
gylcgcgggt atgtgctgc ggggttgcc ggcgggtta ttgctcgtga tgatcgtccg 13740
acagattcca accggaatct ggtggatgcg catcttcac ctccggcgac ttaattttc 13800
gctattctgg agcttgttgt ttatttoggc ctaccgctg ccgggcgggg tcggcgac 13860
ggtaggcgct gtgcagcgc tgatggtgt gttcatctct gccgtctgc taggtagccc 13920
gatacagattg atggcggtcc tgggggtat ttgcggaact gcgggcgtgg cgctgttgtt 13980
gttgacacca aacgcagcgc tagatcctgt cggcgctgca gcgggcctgg cggggcggtt 14040
ttcatcgcg ttccggaacg tgcgacccg caagtggcaa cctcccgtc ctctgctcac 14100
ctttaccgcc tgycaactgg cggccggagg acttctgtc gttccagtag ctttagtgtt 14160
tgatccgcc atcccgatgc ctacaggaac caatgttctc ggctcggtgt ggcctcgct 14220
gacggagcg ggttaacct acttcccttg gttccggggg atctcgcgac tggaaacct 14280
agttgtttcc ttactgggct ttctcagccc cagatctggg gtcatcagc cggggatgca 14340
tcaggccgac agtcggaact tcgggtccc gacctgtacc attcggtgag caatggatag 14400
gggagttgat atogtcaacg ttcaattcta aagaaatagc gccactcagc ttactcagcg 14460
gcttatcca gogatttctt attatgtcgg catagttctc aagatcgaca gcctgtcacg 14520
gttaagcgag aatgaataa gaaggctgat aattgggac tctgcgagg agatgatatt 14580

tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgtaccct ccgcgagatc 14640
atccgtgttt caaacccggc agcttagttg ccgttcttcc gaatagcatc ggtaacatga 14700
gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccgact gatgggctgc 14760
ctgtatcgag tgggtatgtt gtgcgagct gccggtcggg gagctgttgg ctggctgggtg 14820
gcaggatata ttgtgtgtga aacaaattga cgttagaca acttaataac acattcggga 14880
cgtttttaat gtactggggt ggtttttctt ttaccagtg agacgggcaa cagctgattg 14940
cccttcacgg cctggccctg agagagttgc aqcaagcggc ccacgctggt ttgcccagc 15000
aggcgaaaat cctgtttgat ggtggttccg aaatcgcaa aatcccttat aaatcaaaag 15060
aatagccga gataggggtg agtgtgttc cagtttgaa caagagtcca ctattaaaga 15120
acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg 15180
aaccatcacc caaatcaagt tttttggggt cgaggtgccg taaagcacta aatcggaacc 15240
ctaaaggag ccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg 15300
aagggaagaa agcgaaagga gggggcgcca ttacggtgc gcaactgtt ggaaggcgca 15360
tcgggtcggy cctcttcgct attacgcaag ctggcgaaag gggatgtgc tgcaaggcga 15420
ttaagttggg taacgccagg gttttccag tcaagacgtt gtaaaacgac ggccagtga 15480
ttcagctcgt gtaccgggg atctttcgac actgaaatac gtogagcctg ctccgcttgg 15540
aagcggcgag gagcctcgtc ctgtcacaac taccaacatg gactacgata agggccagtt 15600
ccgcagctc attaagagcc agttcatggg cgttgcatg atggcgctca tgcattgtga 15660
cttcaagtag accaacgctc ttctgatcca gtcatcatc cgtgaaggc gctttcgaat 15720
ctggttaaga tccagctctt cgggaagcca gcgactggtt acctccagcg tcccttaag 15780
gctccaaca gctttctcay ccayggccag cccaagacg acaaggctc cctccagaac 15840

gocgagaaga actggagggg tgggtgcaag gaggagtaag ctccttattg aagtcggagg 15900
acggagcggg gtcaagagga tattcttcga ctctgtatta tagataagat gatgaggaat 15960
tggaggtagc atagcttcat ttggatttgc ttccaggct gagactctag cttggagcat 16020
agagggtcct ttggctttca atattotcaa gtatctcgag ttgaaetta ttccctgtga 16080
accttttatt caccaatgag cattggaatg aacatgaatc tgaggactgu aalcgccatg 16140
aggttttoga aatacatccg gatgtogaag gcttggggca cctgcgttgg ttgaatttag 16200
aacgtgggac tattgatcat ccgatagctc tgcaaagggc gttgcacaat gcaagtcaaa 16260
cgttgctago agttccagggt ggaatgttat gatgagcatt gtattaaatc aggagatata 16320
gcatgatctc tagtttagctc accacaaaag tcagacggcg taaccaaag taacacaaca 16380
caagctgtaa ggatttcggc acggctacgg aagacggaga agccacctc agtgagctcg 16440
agtaaccttt aattctattt gtgtttgatc gagacctaat acagccocta caacgacct 16500
caaagtcgta tagctaccag tgaggaagtg gactcaaatc gacttoagca acatctcctg 16560
gataaacctt aagcctaaac tatacagaat aagatagggt gagagcttat accgagctcc 16620
caaatctgtc cagatcatgg ttgacctgtg cctggatctt cctatagaat catccttatt 16680
cgttgacctc gctgattctg gagtgaacca gagggtcatg acttgagcct aaaatccgcc 16740
gcctocacca ttgtagaaa aatgtgacga actcgtgagc totgtacagt gaccggtgac 16800
tctttctggc atgcggagag acggacggac gcagagagaa gggctgagta ataagccact 16860
ggccagacag ctctggcggc tctgaggtgc agtgatgat tattaatccg ggaccggccg 16920
ccctccgcc ccgaagtga aaggtgggtg tgcccctcgt tgaccaagaa tctattgcat 16980
catcgagaa latggagctt catcgaatca ccggcagtaa gcgaaggaga atgtgaagcc 17040
aggggtgtat agcctoggo gaaatagcat gccattaacc taggtacaga agtccaattg 17100

cttcgatctc ggtaaaagat tcacagata gtacctctc cgaagtaggt agagcgagta 17160

cccgcgcggt aagctcccta attggcccat cggcctctg tagggcgctc aaatatcgtg 17220

cctctctgc tttgcccggt gtagaanc ggaaaggccg ctccaggact gccagcggc 17280

gcagaccggg aacacaagct ggcagtcgac ccacccggg ctctgcactc gacctgctga 17340

ggctccctcag tccttgtag gcagctttgc ccgtctgtc cggccgggtg gtccgggggg 17400

ttgacaaggt cgttgctca gtccaaatt tgttgccata tttctctgt ctcccacca 17460

gctgctcttt tctttctct tcttttccc atcttcagta tattcatctt ccacccaag 17520

aacctttatt tccctaagt aagtacttg ctacatccat actccatct tccatccct 17580

tattcctttg aacctttcag ttcgagcttt ccacitlual cgcagcttga ctacageta 17640

ccccgcttga gcagacatca ccagcctga actcaccgc acgtctgtcg agaagttct 17700

gatcgaaaag ttgcacagcg tctccgacct gatgcagctc tcggaggggg aagaatctcg 17760

tgctttcagc ttgatgtag gaggcgctgg atatgtcctg cgggtaata gctgcgcga 17820

tggtttctac aaagatcgtt atgtttatcg gcactttgca tcggccgcgc tccgattcc 17880

ggaagtgcct gacattgggg aattcagcga gacctgacc tattgcactt ccgcctgctc 17940

acagggtgct acgttgcaag acctgcctga aaccgaactg ccgctgttc tgcagccgggt 18000

cgcggaggcc atggatgcga tcgctgcggc cgaicttagc cagacgagcg ggttcggccc 18060

attcggaccg caaggaaatc gtcaatacac tacatggcgt gatttcatat gcgcgattgc 18120

tgatcccat gtgtatcact ggcaactgt gatggagac accgtcagtg cgtccgtgc 18180

gcagctctc gatgagctga tgctttgggc cgaggactgc ccgaagtc gccacctcgt 18240

gcacgggat ttcggctcca acaatgtct gacggacaat gccgcataa cagcggctcat 18300

tgactggagc gaggcgatgt tcggggattc ccaatacgag gtgcccaaca tcttctctg 18360

gaggcgctgg ttgcttgta tggagcagca gacgcgctac ttcgagcgga ggcacccgga 18420
 gcttgacgga tcgucygcgc tccgggcgta tatgctccgc attggtcttg acnaactcta 18480
 tcagagcttg gttgacggca atttcgatga tgcagcttgg gcgcagggtc gatgcgacgc 18540
 aatcgtccga tccggagccg ggactgtcgg gcgtacacaa atcgcccgca gaagcgcgcc 18600
 cgtctggacc gatggtcttg tagaagtact cgcgatagt ggaaaccgac gccccagcac 18660
 tcgtccgagg gcaaaggaat agagtagatg ccgaccgagg gatcgatcca cttaacgtta 18720
 ctgaaatcat caaacagctt gacgaatctg gatataagat cgttggtgtc gatgtcagct 18780
 ccggagttag gacaaatggt gttcaggatc tcgataagat acgttcattt gtccaagcag 18840
 caaagagtgc cttctagtga tttaatagct ccaclylcaac aagaataaaa cgcgttttgc 18900
 ggtttacctc ttccagatgc agctcatctg caatgcatta atgcattgac tgcaacctag 18960
 taacgccttn caggtctcgg cgaagagaag aatagcttag cagagctatt ttcattttgc 19020
 ggagacgaga tcaagcagat caacggtcgt caagagacct acgagactga ggaatccgct 19080
 cttggttcca cgcgactata tatttgtctc taattgtact ttgacatgct cctcttcttt 19140
 actctgatag cttgactatg aaaattccgt caccagcncc tgggttcgca aagataattg 19200
 catgtttctt ccttgaactc tcaagcctac aggacacaca ttcacgttag gtataaacct 19260
 cgaaatcant tctactaag atggtatata atagtaacca tgcattygly cctagtgaat 19320
 gctccgtaac acccaatacg cgggcgaaa cttttttaca actctctat gagtcgttta 19380
 cccaqaatgc acaggtacac ttgtttagag gtaactcttc tttctagcta gaagtccctg 19440
 tgtactgtgt aagcgccac tccacatctc cactcgacct gcaggcatgc a 19491

<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (3471)..(3471)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3679)..(3679)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3770)..(3770)
<223> n is a, c, g, or t

<400> 46
gatcttttoga caactgaaata cgtcgagcct gctcgccttg gaagcggcga ggagcctcgt 60
cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120
cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180
cttctgatcc agtcgatcat ccgctgaagg cgcttttogaa tctggttaag atccacgtct 240
tcgggaagcc agcgactggt gacctccagc gtccctttta yyctgocac agcttttctca 300
gccaggggcca gcccaagacc gacaaggcct cctccagaa cgcggagaag aactggaggg 360
gtggtctcaa ggaggagtaa gctccttatt gaagtoggag gaaggagcgg tgtcaaggag 420
atattcttgc actctgtatt atagataaga tgataggaa ttggaggtag catagcttca 480
tttggatttg ctttccaggc tgagactcta gcttggagca tagagggtcc tttggcttcc 540
aatattctca agtatctoga gtttgaactt attccctgtg aaccttttat tcaccaatga 600

gcattggaat gaacatgaat ctgaggactg caatgccat gaggttttcg aaatacatcc	660
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca	720
tccgatagct ctgcaaaggc cgttgacaaa tgcaagtcaa acgttgctag cagttccagg	780
tggaaatgta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct	840
caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg	900
cacggctacg gaagacggag aagocacctt cagtggactc gagtaccatt taattctatt	960
tgtgtttgat cgagacctaa tacagccctt acaacgacca tcaaagtcgt atagctacca	1020
gtgaggaagt ggactcaaat cgacttcagc aacatctoct ggataaactt taagcctaaa	1080
ctatacagaa taagataggt ggagagctta taacyagctc ccaaatotgt ccagatcatg	1140
gttgaccggt gcttggtct tctatagaa tcatccttat tcttgacct agctgattct	1200
ggagtgaccg agagggtcat gacttgagcc taaaatccgc cgcctccacc atttgtagaa	1260
aaatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttctgg catgcgagga	1320
gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg	1380
ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc ccggaagtgg	1440
aaaggctggt gtgccctcgt ttgaccaaga atctattgca tcatcgga atattggagct	1500
tcatcgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg	1560
cgaaatagca tgccattaac ctaggtagag aagccaatt gcttcgagc ttgtaaaaga	1620
ttcacgagat aqtaccttct ccgaagttag tagagcgagt acccgggcgg taagctccct	1680
aattggccca tcggcatct gtagggcgtc caaatatcgt gcctctcctg ctttgcccg	1740
tgatgaaac cggaaaggcc gtcaggagc tggccagcgg cgcagaccgg gaacacaagc	1800
tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtcctcgga	1860

ggcagctttg ccccgctctg cgcgccggtg tgcggcgagg gttgacaagg tcgttgcgtc 1920
 agtccacat ttgttgcat atlllcclyc tctccccc agctgctctt tctttttctc 1980
 ttttttttc catcttcagt atattcatct tcccatccaa gaacctttat tcccttaag 2040
 taagtacttt gctacatcca tactccatcc tcccatccc ttattctttt gaacctttca 2100
 gttcgagctt tccacttca tgcgagcttg actaacagct acccgcttg agcagacatc 2160
 accatgcttg aactcacgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220
 gtctccgacc tgatgcagct ctgggagggc gaagaatctc gtgctttcag ctctgatgta 2280
 ggagggcggt gatatgtctc gggggtaaat agctgcgcgc atggtttcta caaagatgt 2340
 tatgtttatc ggcactttgc atcgccgcgc ctcccgattc cygaaglyct tgacattggg 2400
 gaattcagcg agagctgcac ctattgcato tcccgccgtg cacaggggtg caggttgcaa 2460
 gacctgcctg aaaccgaact gccgcgtgtt ctgcagccgg tcgggagggc catggatgag 2520
 atcgtcgagg cagatcttag ccagacgagc gggttcggcc cattcggacc gcaaggatc 2580
 ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640
 tggcaaatcg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700
 atgctttggg ccgaggactg ccccgaaatc cggcacctcg tgcacgggga tttcggctcc 2760
 aacaatgtcc tgacggacaa tggccgcata acagcggta ttactggag cagggcgaly 2820
 ttccggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880
 atggagcagc agacccgcta ctccgagcg aggcatccgg agcttcgagg atcccgcggg 2940
 ctccggcgct atatgctccg cattggtctt gaccaactct atcagagctt ggttgacggc 3000
 aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgctcg atccggagcc 3060
 gggactgtcg ggcgtacaca aatcgccgc agaagcgagg ccgtctggac cgatggctgt 3120

gtagaagtac tcgccgatag tggaaaccga cgccccagca ctogtccgag ggcaaaggaa 3180
 tagagtagat gccgcgcgog ggaatcaton anttaacgtt actgaatatca tcaaacagct 3240
 tgacgaatct ggaataaaga togttgggtg cgatgtcagc tcoggagttg agacaaatgg 3300
 tgttcaggat ctgcataaga tacgttcatt tgtccaagca gcaaagagtg cttcttagtg 3360
 atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420
 cagctcatct gcaatgcatt aatgcattga ctgcaacct gtaacgcctt ncaggtcccg 3480
 gcgaagagaa gaatagetta gcagagctat ttctattttc gggagacgag atcaagcaga 3540
 tcaacggctg tcaagagacc tacgagactg aggaatccgc tcttggctcc acgcgactat 3600
 atatttctct ctaattylau ttgacatgc tectettott taectcgata gcttgcattat 3660
 gaaaattcog tcaccagcnc ctgggttcgc aaagataatt gcattgtttc tcttgaact 3720
 ctcaagccta caggacacac attcatcgta ggtataaacc togaatacan ttectactaa 3780
 gatgggtatac aatagtaacc atgcatggtt gcctagttaa tgctccgtaa caccoatac 3840
 googgoogaa acttttttac aactctocta tgagtogttt acccagaatg cacaggtaaa 3900
 cttgtttaga ggtaatcctt ctttttagct agaagtcctc gtgtactgtg taagcgccca 3960
 ctccacatct ccactogacc tgcaggcatg caagcttgaa ttogagctcg gtaccggggg 4020
 atctttcgac actgaataac gtccagcctg ctccgottgg aagcggcgag gagcctcgtc 4080
 ctgtcacaac taccaacatg gattacgata agggccagtt ccgcacgctc attaagagcc 4140
 agttcatggg cggtggcatg atggccgtca tgcattctga cttcaagtac accaacgctc 4200
 ttctgatcca gtccatcctc cgctgaaggc gctttcgaat ctggttaaga tccacgtctt 4260
 cyggaagcca gogactggtg aactccagcg tccctttaag cctgccaaca gctttctcag 4320

ccagggccag cccaagaccg acaaggccctc cctccagaac gccgagaaga actggaggggg 4380
 tgggtgcaag gaggagtaag ctcttattg aagtcggagg acggagcggg gtcaagagga 4440
 tattcttcga ctctgtatta tagataagat gatgaggaat tggaggtagc atagcttcat 4500
 ttggatttgc tttccaggct gagactctag cttggagcat agagggctct ttggctttca 4560
 atattctcaa gtatctcgag tttagaacta ttccctgtga accttttall caccantgag 4620
 cattggaatg aacatgaatc tgaggactgc aatcgccatg aggttttcga aatacatccg 4680
 gatgtogaag gnttggggca cctgcgttgg ttgaatttag aacgtggcac tattgatcat 4740
 ccgatagctc tgcaaaagggc gttgcacaat gcaagtcaaa cgttgctagc agttccaggc 4800
 ggaatgttat gatgagcatt gtattaaatc aggagatata gcatgatctc tagtttagctc 4860
 accacaaaag tcagacggcg taacaaaag tcacacaaca caagctgtaa ggatttcggc 4920
 acggtactcg aagacggaga agcoacttc agtggaactc agtaccattt aattctattt 4980
 gtgtttgac gagacctaat acagccocta caacgacat caaagctgta tagctacua 5040
 tgaggaagtg gactcaaatc gaattcagca acatctctg gataaacttt aagcctaacc 5100
 tatccagact aagataggcg gagagcttat accgagctcc caaatctgct cagatcatgg 5160
 ttgacoggcg cctggatctt octatagaat catccttatt cgttgaccta gctgattctg 5220
 gaggtagcca gagggctatg acttgagcct aaatccgcc gctccacca tttgtagaaa 5280
 aatgtgacga actcgtgagc tctgtacagt gaccggtgac tctttctggc atgaggagag 5340
 accgacggac gcagagagaa gggctgagta ataagccact ggccagacag ctctggcgcc 5400
 tctgaggtgc agtgatgat tattaatccg ggaacggccg cccctccgcc cgaagtggga 5460
 aaggtcgtg tgccctcgt tgaccaagaa tctattgcat catcgagaaa tatggagctt 5520
 catogaatca ccggcagtaa ggaaggaga atgtgaagcc aggggtgtat agccgtcgcc 5580

gaaatagcat gccattaacc taggtacaga agtccaattg ctccgatct ggtaaaagat 5640
 tcacagagata gtacctcttc cgaagtaggt agagcgagta cccggcgctg aagctcccta 5700
 attggcccat ccggcatctg tagggcgctc aaatatcgtg cctctcctgc ttggcccggt 5760
 gtatgaaacc ggaaggccg ctcaggagct ggccagcggc gcagaccggg aacacaagct 5820
 ggcagtgcac ccataccggtg ctctgcactc gaactgtga ggtcculcag tccctggtag 5880
 gcagctttgc ccgctctgtc cgcocgggtg gtcggcgggg ttgacaaggt cgttgctgca 5940
 gtccaaacatt tgttgacata ttttctgct ctcccacca gctgctcttt tctttctct 6000
 ttcttttccc atcttcagta tattcatctt cccatccaag aacctttatt tcccctaagt 6060
 aagtaacttg ctacatccat actccatctt tccatccct tattccttg aacctttcag 6120
 ttcgagcttt cccacttcat cgcagcttga ctaacagcta cccogctga gcagacatca 6180
 ccatgtcaat actcacttat ctggaatttc atctctacta tacactacct gtccctggcg 6240
 cattgtgtg gctgctaaag ccgtttcact cacagcaaga caatctcaag tataaatltt 6300
 taatgttgat ggcgcctct accgcatga ttgggacaa ttatatcgtt tatcatcgcg 6360
 cttgylgta ctgtctact tgtgttgtg ctgtcattgg ctatgtacct ctagaagaat 6420
 acatgttctt tatoatcatg actttaatga ctgtcgcgtt ctcaaaacttt gttatgcgtt 6480
 ggcacttgca tactttcttt attagaccca acactctctg gaagcaaa caattagtac 6540
 gccttggtgc tgtttcagct ttattggcaa tcacttatca tgcctggcac ttgacactgc 6600
 caaataaacc ttcattttat ggttcatgca tcttttgta tgttggtctt gtgttggtca 6660
 ttctttggct ggggtgctgc gaatatatct tgcgtcgacc tgtggctgct cttttgtcta 6720
 ttgttatccc tagtgtatc ctatgttggg ctgatatcgt cgtattagt gctggcacat 6780
 ggcataattc tctlgacaa agcactggca aatggtaagt accgatatta cctgtagaag 6840

aatgcctggt ttttaactttg atcaacacag tcttggtttt tgctacctgt gctatagacc 6900

gcgctcaggc catctccat gtgagcgcg gtaatacgac tcaatatagg gcgaattgga 6960

gtccacgcg ggtggcgccg gctctagaac tagtggatcc cccgggctgc aggaattcgg 7020

cacgagctac atttacaag cccgtgagcg gtgcaagcgc tctgcccac atcgcccac 7080

ctctcatct ccactcggtca ttgctgcta ccacgatgct gtggaagctg cagtcaatca 7140

gcgtcaaggc ccgcgcgctt gaactagccc gcgacatcac gcggcccaaa gtctgcctgc 7200

atgctcagcg gtgctcgcta gttcggctgc gagtggcagc accacagaca gaggaggcgc 7260

tgggaacgct gcaggctgcc ggcgcggggc atgagcacag cgcgatgta gcactccagc 7320

agcttgaccg ggtcatcgca gagctcgtg cccggcgcaa acgggagcag ctgtcatacc 7380

aggctgcgcg cattgcagca tcaattggcg tgtcaggcat tgccatcttc gccactacc 7440

tgagatttgc catgcacatg accgtggggc gcgcagtgcc atgggggtgaa gtggctggca 7500

ctctctctt ggtgggttgt ggcgcgctcg gcattggagat gtatgccgcg tatgcacaca 7560

aagccatctg gcattgagtc cctctgggct ggctgctgca caagagccac cacacacctc 7620

gcactggacc ctttgaagcc aacgacttgt ttgcaatcat caatggactg cccgccatgc 7680

tcctgtgtac ctttggcttc tggctgccc acgtcctggg ggcggcctgc tttggagcgg 7740

ggctgggcat cagctatcc ggcattggat atatgtttgt acacgatggc ctgggtgcaca 7800

ggcgctttcc caccgggccc atcgctggcc tgcctacat gaagcgccctg acagtggccc 7860

accagctaca ccacagcgcg aagtaaggct gcgcgcctg ggttatgttc ttgggtccac 7920

aggagctgca gcacattcca ggtcgggcgg agggagtgga gcgactggtc ctgggaactgg 7980

actggtccaa gcgggctcag gccatcctcc atctgtacaa atcatctgtt caaaatcaaa 8040

accctaaaca agccatttcc cttttccagc atgtcaaaga gctagcatgg gccttctgtc 8100

ttcctgacca aatgctaac aatgaattgt ttgatgatct tactatcagc tgggatattt 8160
tactgtaaagc ctcaagaatga ttctatactg catctgcctt ttttccaagt tatgtacgtc 8220
aagacttggg tgttctctat gctttctgca gagctaccga tgacctgtgc gatgtgaat 8280
ccaaatctgt tcaagaaga agagaccaat tagatcttao tcgacaattt gttcgtgac 8340
tcttttagca aaagaccagt ggcctatttg tgattgattg ggaattgtat caaaaccaac 8400
ttcctgcttc ttgtatatca gcccttagag cctttactcg ccttcgccat gtccttgaag 8460
tagaccctgt agaagaacta ttagatgggt acaaatggga tcttgagcgt cgtcctatcc 8520
ttgatgaaca agacttggag gcatactctg cttgtgtggc cagttagtgt ggtgaatgt 8580
gcacacgtgt gattcttgct caagaccaa agyaaaatga tgcctggata attgacctg 8640
cacgtgagat ggggctgggt ctacaatacg ttaacattgc tcgagacatt gtgactgata 8700
gcgagactct gggctcagtg tatctgcctc aacaatggct tagaaaagaa gaaacagaac 8760
aaatacagca aggcacagcc cgtagcctag gtgatcaag actgttgggc ttgtctctga 8820
agctttagg aaaggcagac gctatcatgg tgagagctaa gaagggcatt gacaagttgc 8880
cggcaaatct tcaaggcgggt gtacgagctg cttgccaaagt atatgctgca attggatctg 8940
tactcaagca gcagaagaca acatatccta caagagctca tctaaaagga agcgaaogtg 9000
ccaagattgc tctgttgagt gtatacaacc tctatcaatc tgaagacag cctgtggctc 9060
tccgtcaagc tagaaagatt aagagttttt ttgttgatta gtgaattttt gttttattta 9120
tgtctgataq ttcaataaag agacaacaca tacaataaa aatcattgtc tttaaattgt 9180
aatttagtag agtgtaaagc ctgcattttt ttgttacgca taaacaatga gttcaccccg 9240
cttctgggtt ttaaataatt atgtcaaac agggaaaatt cttttttttc tctctgtctc 9300
ttttttggct tgttgtggag tcacaggctt gtcttcagat tgatagaggt tgtatacact 9360

caacagagca atcttggcac gtctgcttcc ttttagatga gctctttag gatatgttg 9420
cttctgctgc ttgagtacag atccaattgc agcatataat tggcaagcag ctctacaccc 9480
gccttgacag tttgcggca acttgatcaat gcccttcta gctctcacca tgatagcgtc 9540
tgctttct acaagcttgg cgtaatcatg gtcatactg tttctgtgt gaaattgtta 9600
tccgtcaca attccacaca acatacgagc cggaagcata aagtgtaaag cctgggggtc 9660
ctaagtgtg agctaactca cattaattgc gttgctca ctgccgctt tccagtcggg 9720
aaacctgtgc tgccagctgc attaatgaat cgccaacgc gcggggagag gcggtttgcg 9780
tattgggcca aagacaaaag ggcgacattc aaccgattga gggagggaag gtaaatattg 9840
acggaaatta ttoattaaag gtgaattatc accgcaaccg acttgagcaa tttgggaatt 9900
agagccagca aaatcaccag tagcaccatt accattagca aggcgggaaa cgtcaccaat 9960
aaaaccatcg atagcagcac cgtaatcagt agcgacagaa toaagtttgc ctttagcgtc 10020
agactgtagc gcgttttcat cggcatttcc ggtcatagcc ccttatttag cgtttgccat 10080
cttttataa tcaaaatcac cggaaccaga gccaccaccg gaaccgcctc cctcagagcc 10140
gccacctca gaaccggcac cctcagagcc accacctca gagccggcac cagaaccacc 10200
accagagccg ccgccagcat tgacaggagc ccgatctag taacatagat gacacggcg 10260
gcgataattt atcctagttt gcgcgtata ttttgttttc tatccylat taattgtata 10320
attcgggac totaatcata aaaaccatc toataataa cgtcatgcat tacatgttaa 10380
ttattacatg cttaacgtaa ttcaacagaa attatatgat aatcatcgca agacgggcaa 10440
caggattcaa tottaagaaa ctttattgcc aaatgtttga acgatgggg atcatccggg 10500
tctgtgggg gaactccagc aaaatatccg aacgcagcaa gatatcgcg tgcatctcgg 10560
tcttgctgg gcagtcccg ccgacgccgt tgatgtggac gcggggcccg atcatattgt 10620

cgctcaggat cgtggcgcttg tgcttgctcg cgttgctgt cgtaatgata tcggcacctt 10680
cgaccgcttg ttccgcagag atcccgctgg cgaagaactc cagcatgaga tccccgcgt 10740
ggaggatcat ccagccggcg tcccgaaaa cgattccgaa gcccaacctt tcatagaag 10800
cggcggtgga atcgaaatct cgtgatggca ggttggcggt cgcttggtcg gtcatttoga 10860
accccagagt cccgctcaga agaactcgtc aagaaggcga tagaaggcga tgcgtcgca 10920
atcgggagcg gcgataccgt aaagcacgag gaagcggtca gccattcgc cgccaagtc 10980
ttcagcaata tcacgggtag ccaacgctat gtctgatag cggtcgccca caccagccg 11040
gccacagtgc atgaatccag aaaagcgcc attttccacc atgatattcg gcaagcaggc 11100
atgcccatgg gtcacgacga gatcatcgcc gtggggcatg cggcgcttga gcttggcga 11160
cagttcggtt ggcgcgagcc cctgatgctc ttcttcaga tcctctgat cgaacagacc 11220
ggcttcacat cgagtaoctg ctgcctgat gcgatgttc gcttggtggt cgaatgggca 11280
ggtagccgga tcaagcgtat gcagccgcc cattgcatca gccatgatgg atactttctc 11340
ggcaggagca aggtgagatg acaggagatc ctgccccggc acttcgcca atagcagcca 11400
gtccctccc gcttcagtga caacgtcgag cacagctgcg caaggaaagg ccgtcgtggc 11460
cagccacgat agcgcgcgtg cctcgtctg cagttcattc agggcacagg acaggctgg 11520
cttgacaaaa agaaccgggc gccctgcgc tgacagccgg aacacggcg catcagagca 11580
gcgatgtgc tgttgtgccc agtcatagcc gaatagcctc tccaccaagg cggccggaga 11640
acctcgctgc aatccattt gttaaatcat gcgaacgat ccagatccgg tgcagattat 11700
ttggattgag agtgaatat agactctaatt tggataccga ggggaattta tggaaagtca 11760
gtggagcatt ttgaaaga aatatttgt agctgatagt gaccttaggc gacttttgaa 11820

cgcgcaataa tggttttctga cgtatgtgct tagctcatta aactccagaa acccgcggtc 11880
gagtggctcc ttcaacgttg cggttctgtc agttccaaac gtaaacggc ttgtcccgcg 11940
tcactggcgg ggggtcataac gtgactccct taattctccg ctcatgatca gattgtcgtt 12000
tcccgccttc agttttaaact atcagtgttt gacaggatat attggcgggt aaacctaaaga 12060
gaaaagagcg tttattagaa taatcggata tttaaaaggc cgtgaaaagg tttatccgtt 12120
cgtccatttg tatgtgcatg ccaaccacag ggttcccag atctggcgcc ggccagcgag 12180
acgagcaaga ttggcgcgcg cccgaaacga tccgacagcg cgcccagcac aggtgcgcag 12240
gcaaattgca ccaacgcata cagcgccagc agaatgccat agtggcggtt gacgtcgttc 12300
gagtgaacca gatcgcgag gagggccggc agcaccggca taatcaggcc gatccgcaca 12360
gcgtcgagcg cgacagtgtc cagaattacg atcaggggta tgttgggtt cactctggc 12420
ctccggacca gccctcgtg gtccgattga acgcgcggtt tctttatcac tgataagttg 12480
gtggacatat tatgtttatc agtgataaag tgtoaagcat gacaaagtgt cagccgaata 12540
cagtgatccg tgcgcgcctg gacctgttga acgaggtcgg cgtagacggt ctgacgacac 12600
gcaaacctggc ggaacgggtt ggggttgaac agccggcgct ttactggcac ttcaggaaaca 12660
agcggggcgt gctcgacgca ctggccgaag ccatgctggc ggagaatcat acgcattcgg 12720
tgccgagagc cgacgacgac tggcgctcat ttctgatcgg gaatgccgc agcttcaggc 12780
aggcgctgct cgcctaccgc gatggcgcgc gcattccatgc cggcagcgca ccggggcgac 12840
cgcagatgga aacggcgac gcgcagcttc gcttccctcg cagggcggtt ttttcggcg 12900
gggacgcgt caatgcgctg atgacaatca gctacttcaac tgttggggcc gtgcttgagg 12960
agcagggcgg cgacagcgat gccggcgagc gggcgggcac cgttgaaacg gctccgctct 13020
cgccgctgtt gggggccgcy alagacgcct tggaggaagc cggtcgggac gtagcgttcg 13080

agcagggact cgcgggtgatt gtcgatggat tggcgaaaag gaggctcggt gtcagggaacg 13140

ttgaaggacc gagaagggt gacgattgat caggaccgct gccggagcgc aaccacactca 13200

ctacagcaga gccatgtaga caacatcccc tcccccttcc caccgcgtca gacgcccgta 13260

gcagcccgct acgggctttt tcatgccctg ccttagcgtc caagcctcac ggccgcgctc 13320

ggcctctctg ggcgccttct ggcgctcttc cgcttcctcg ctcaactgact cgctgcgctc 13380

ggctggtcgg ctgcggcgag cggtatcagc tcaactcaaag gcggtaatac ggttatccac 13440

agaatcacgc gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggcaggaa 13500

ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg acgagcatca 13560

caaaaatoga cgctcaagtc agagggtggcg aaaccgcaca ggactataaa gataccaggc 13620

gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgcgcg ttaccggata 13680

cctgtccgcc tttctccctt cgggaagcgt ggcgcttctc cgctgcataa ccctgcctcg 13740

gggtcattat agcgattttt tcggtatata catccttttt cgcacgatat acaggatttt 13800

gccaaagggt tcgtgtagac ttctcttggt gtatccaacg gcgtcagccg gccaggatag 13860

gtgaagtagg cccaccgcg agcgggtggt ccttcttcac tctcccttat tcgcacctgg 13920

cgggtgctcaa cgggaatcct gctctgcgag gctggccggc taccgcggcg gtaacagatg 13980

agggcaagcg gatggctgat gaaaccaagc caaccaggaa gggcagccca cctatcaagg 14040

tgtactgcct tccagacgaa cgaagagcga ttgaggaaaa ggccggcgcg gccggcatga 14100

gcctgtcgcc ctacctgctg gccgtcgcc agggctacaa aatcacggcg gtcgtggact 14160

atgagcacgt ccgcgagctg gcccgcatca atggcgacct gggccgctcg ggccgcctgc 14220

tgaactctg gctcacgcac gaccgcgca cggcgcggtt cgggtgatgcc acgatcctcg 14280

ccctgctgca gaagatcgaa gagaagcagg acgagcttgg caaggtcatg atggcgctg 14340

tccgcgccgag ggcagagcca tgactttttt agccgctaaa acggccgggg ggtgcgcgtg 14400

attgccaagc acgtcccat gcgtccatc aagaagagcg acttcgcgga gctggtgaag 14460

tacatcacgc acgagcaagg caagaccgag cgcctttgag acgctcacgc ggttggttgc 14520

cctgcgcgtc gggctggcgg ccgtctatgg ccctgcaaac gcgccagaaa cgcgctcgaa 14580

gcctgctgag ayacaccgc gccgcggcgg ttgtggalac ctgcggaaa acilggccct 14640

cactgacaga tgagggggcg acgttgacac ttgaggggac gactcacccg gcgcggcggt 14700

gacagatgag gggcagctc gatttcggcc ggcgacgtgg agctggccag cctcgcaaat 14760

cgcgcaaac gcctgatttt acgcgagttt cccacagatg atgtggacaa gcctggggat 14820

aagtgcctg cgggtattgac acttgagggg cgcgactact gacagatgag gggcgcgatc 14880

cttgacactt ggggggcaga gtgctgacag atgagggggc cacctattga catttgaggg 14940

gctgtccaca ggcagaaaaa ccagcatttg caagggttcc cgcgcgtttt tcggccaccg 15000

ctaacctgtc ttttaacctg cttttaaacc aatatttata aacctgtttt ttaaccaggg 15060

ctgcgcctg tgccgtgac cgcgcacgac gaagggggtt gcccccctt ctgcaacct 15120

ccggggccgc taacgggggc ctccatccc cccaggggct gcgcctctcg gcgcgaacg 15180

gcctcacccc aaaaatggca gcgctggcag tccttgccat tgccgggacg ggggcagtaa 15240

gggatgggc gatcagccc agcgcgcgc ccggaagcat tgacgtgccc caggtgctgg 15300

catcgacatt cagcgaccag gtgccgggca gtgagggcgg cggcctgggt ggcgcctgcg 15360

ccttacttc gccgctggg gcattcacgg acttcattgg ggggcccgca atttttacct 15420

tgggcattct tggcatagtg gtgcgggtg ccgtgctcgt gttcgggggt gcgataaac 15480

cagcgaacca tttgagtgta taggtaagat tataccaggg tatgaaaacg agaattggac 15540

cttacagaa ttactctatg aagcgccata tttaaaaagc taccaagacg aagaggatga 15600

agaggatgag gaggcagatt gccctgaata tattgacaat actgataaga taatatatct 15660
tttatataga agatatcgcc gtatgtaagg atttcaggag gcaaggcata ggagcgcgcc 15720
ttatcaatat atctatagaa tgggcaaagc ataaaaactt gcattggacta atgcttgaaa 15780
cccaggacaa taaccttata gcttgtaaat tctatcataa ttgggtaatg actccaactt 15840
attgatagtg ttttatgttc agataatgcc cgatgacttt gtcatgcagc tccaccgatt 15900
ttgagaacga cagcgacttc cgtccagacc gtgccaggtg ctgcctcaga ttcagggttat 15960
gcgcctcaat tcgctcgta tatcgcttgc tgattacgtg cagctttccc ttcaggcggg 16020
attcatacag cggccagcca tccgtcatcc atatcaccac gtaaaagggt gacagcagcc 16080
tcataagacg ccccgagctc gccatagtg gttcacgaa tactgcgga aaaaacgtct 16140
tcgggagact gtcatacgcc taaaacagcc agcgcgtggc cgatttagcc ccgacatagc 16200
cccactgttc gtccatttcc gcgcagacga tgacgtcact gcccggtgt atgcgcgagg 16260
ttaccgacty cggcctgagt tttttaagt acgtaaaac gtgttgaggc caacgcccat 16320
aatgcgggct gttgcccgcc atccaacgcc attcatggcc atatcaatga ttttctgggt 16380
cgtaccgggt tgagaagcgg tgtaagtga ctgcagttgc catgttttac ggcagtgaga 16440
gcagagatag cgctgatgtc cggcggtgct tttgccgta cgcaccacc cgctcagtagc 16500
tgaaacaggag ggacagctga tggacacaga agccactgga gccctcctaaa aacaccatca 16560
tactataat cagtaagttg gcagcatcac ccataattgt ggtttcaaaa tcggctccgt 16620
cgatactatg ttatcgcga actttgaaa caactttgaa aaagctgttt tctggtatctt 16680
aagggttttag aatgcaagga acagtgaatt ggagttcgtc ttgttataat tagcttcttg 16740
gggtatcttt aaactactga gaaaagagga aggaaataat aaatggctaa aatgagaata 16800
tcaccggaat tgaaaaaact gatcgaaaa taccgctgcg taaaagatag ggaagggaatg 16860

tctcctgcta aggtatataa gctggaggga gaaaatgaaa acctatattt aaaaatgacg 16920

gacagcccggt ataaaggac cacctatgat gtggaacggg aaaaggacat gatgctatgg 16980

ctggaaggaa agctgcctgt tccaaaggto ctgcactttg aacggcatga tggctggagc 17040

aatctgctca tgagtgggc cgatgggctc ctttgcctgg aagagtatga agatgaacaa 17100

agccctgaaa agattatcga gctgtatgcg gagtgcacga ggctctttca ctccatcgac 17160

atatcggtatt gtccctatac gaatagctta gacagccgct tagccgaatt ggattactta 17220

ctgaataacg atctggccga tgtggattgc gaaaactggg aagaagacac tccatttaaa 17280

gatccgcgcg agctgtatga ttttttaaa acggaagcgc ccgaagagga acttgtcttt 17340

tcccaaggcg acctgggagc nagnaacatc tttgtqaaag atggcaaggt aagtggtttt 17400

attgatcttg ggagaagcgc caggcgagac aagtgggatg acattgcctt ctgcgtccgg 17460

togatcaggg aggatatcgc ggaagaacag tatgtcagac tattttttga ctactgggg 17520

atcaagcctg attggggaga aataaaatat tatattttac tggatgaatt gttttagta 17580

ctagatgtgg cgcaacgatg cggcgacaaa gcaggagcgc accgacttct tccgatcaa 17640

gtgtttttgc tctcaggcgc agggccacgc caagtatttg ggcaagggtt cgttggtatt 17700

cgtgcagggc aagattcga ataccaagta cgagaaggac ggccagacgc tctacgggac 17760

cyacttcatt gcgataagg tggattatct ggacannaa gcaccagcgc ggtcaaatca 17820

ggaataaggg cacattgccc cggcgtgagt cggggcaatc ccgcaaggag ggtgaatga 17880

toggacgttt gaccggaagg catacaggca agaactgac gacgcgggtt tttccgccga 17940

ggatgccgaa accatcgcaa gccgcaccgt catgctgctg ccccgcgaaa ccttccagtc 18000

cgtcggctcg atggtccagc aagctacgc caagatcgag cgcgacagc tgcaactggc 18060

tcccctgcc ctgccgcgc catcgccgc cgtggagcgt tgcgtcgtc tgaacagga 18120

ggcggcaggt ttggogaagt cgatgaccat cgacacgcga ggaactatga cgaccaagaa 18180
gcgaaaaacc gccggcgagg acctggcaaa acagggtcagc gaggncaagc aggnccgctt 18240
gctgaaacac acgaagcagc agatcaagga aatgcagctt tccttggttcg atattggccc 18300
gtggccggac acgatgcgag cgatgcacaa cgacacggcc cgtctgtccc tgttaccac 18360
ggcgaacaag aaaatccgcg gcgaggcgct gcaaaacaag gtcattttcc acgtcaacaa 18420
ggacgtgaag atcacctaca ccggcgtcga gctgcgggcc gacgatgagc aactgggtgtg 18480
gcgacagggtg ttggagtacg cgaagcgcac ccctatcggc gagccgatca ccttcacgtt 18540
ctacagagctt tgcaggacc tgggctggtc gatcaatggc cggtattaca cgaaggccga 18600
ggaatgcctg tcgcccctac aggcgacggc gatgggcttc acgtccgacc gsgttggga 18660
cctggaatcg gtgtcgtgc tgcaccgctt ccgcttctg gaccgtggca agaaaacgtc 18720
ccgttgccag gtctgatcg acgaggaaat cgtgtgtctg tttgtggcg accactaac 18780
gaaattcata tgggagaagt accgcaagct gtgcgcagc gcccgacgga tgttcgacta 18840
tttcagctcg caccgggagc cgtaccgct caagctggaa accttccgcc tcattgtcgg 18900
atcggtattcc acccgctga agaagtggcg cgagcaggtc ggcaagcct gcgaagagtt 18960
gcgaggcagc ggcctgggtg aacacgcctg ggtcaatgat gacctgggtc attgcaaacg 19020
ctagggcctt gtggggtcag ttcgggttg ggggttcagca gccagcgctt taclggcatt 19080
tcaggaaaca gcgggacctg ctgcacgcac ttgcttcgct cagtatcgct cgggacgcac 19140
ggcgcgctct acgaactgcc gataaacaga gattaaat tgacaattgt gattaaggct 19200
cagattcgac ggcttggagc ggccagcgtg caggatttcc gcgagatccg attgtcggcc 19260
ctgaagaaag ctccagagat gttcgggtcc gtttaacgagc acgaggagaa aaagcccatg 19320

gaggcggttcg ctgaacgggtt gcgagatgcc gtggcattcg gcgcctacat cgacggcgag 19380
atcattgggc tgtcgttctt caaacaggag gacggccccg aggaagctca caaggcgcat 19440
ctgtccggcg ttttcgtgga gcccgaaacg cgaggccgag gggtccggcg tatgctgttg 19500
gggcggttcg cggcggggtt attgctcgtg atgatcgcc gacagattcc aacgggaatc 19560
tggtggatgc gcatcttcat cctcggcgca cttatattt cgtattctg gactgttg 19620
tttatttcg tctaccgctt gcggggcggt gtccggcgca cggtaggcgc tgtgcagccg 19680
ctgatggtcg tggtcatctc tgcgcctctg ctaggtagcc cgatacgatt gatggcggtc 19740
ctgggggcta tttcggaac tgcggcggtg gcgctgttg tgttgacacc aaacgcagcg 19800
ctagatcctg tcggcgctgc agcgggcctg gggggggcg tttccatggc gttcgaacc 19860
gtgctgaccc gcaagtgga cctcccggt cctctgctca ctttacgcg cttgcaactg 19920
ggggcgag gacttctgct cgttcagta gcttagtgt ttgatccgc aatcccgatg 19980
cctacaggaa ccaatgttct cggcctggcg tggctcgcc tgatcggagc ggylltaacc 20040
tacttcttt ggttcgggg gatctcgca ctgaaacta cagttgttc cttactggcg 20100
tttctcagcc ccagatctgg ggtcgatcag cgggggatgc atcaggccga cagtcggaac 20160
ttcgggtccc cgacctgtac cattcggtga gcaatggata ggggagtga tatcgtcaac 20220
gttcacttct aaagaaatag cgcactcag cttctcagc ggctttatcc agcgatttcc 20280
tattatgtcg gcatagtctt caagatcgac agcctgtcac ggtaagcga gaaatgaata 20340
agaaggctga taattcggat ctctcggag gagatgatat ttgatcacg gcagcaacgc 20400
tgtgtcatcg ttacaatcaa catgctaccc tcgcgagat catcgtgtt tcaaacccgg 20460
cagcttagtt gccgttcttc cgaatagcat cggtaacatg agcaaatgtt gccgccttac 20520
aacgcctctc cgcgtgacgc cgtcccgac tgatgggctg cctgtatcga gtgggtgattt 20580

tgtgccgagc tgcgggtcgg ggaagctgtt gctggctggg ggcaggatat atttggtgt 20640
aaacaaattg acgcttagac aacttaataa cacattgcgg acgtttttaa tgtactgggg 20700
tggtttttct tttaccagt gagcgggca acagctgatt gcccttcacc gectggccct 20760
gagagagttg cagcaagcgg tcacgctgg tttgccccag caggcgaaaa tectgtttga 20820
tggtggttcc gaaatcggca aaatccctta taaatcaaaa gaatagcccg agatagggtt 20880
gagtgttgtt ccagtttga acaagagtc actattaaag aacgtggact ccaacgtcaa 20940
agggcgaaaa accgtctatc agggcgatgg cccactacgt gaaccatcac ccaaatcaag 21000
ttttttgggg tcgaggtgcc gtaaagcaat aaatcggaac cctaaagggg gcccccgatt 21060
tagagcttga cggggaaagc cggcgaaagt ggcgagaaa gaaaggaaga aagcgaaagg 21120
agcggggccc attcaggctg cgcaactgtt gggaaggcgg atcggtgcgg gcctcttcgc 21180
tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag 21240
ggttttccca gtcacgacgt tgtaaacga cggccagtga attcgagctc ggtaccggg 21300

<210> 47

<211> 17756

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n i s a , c , g , o r t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n i s a , c , g , o r t

<400> 47

ccgggctgggt tgcctcgcgc gctgggctgg cyggcgtcta tggcctgca aacggccag 60

aaacgcctc gaagcgtgt gcgagacacc gggccgccg gcgttgtgga tacctgcgg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcggcgc gttgacagat gaggggcagg ctgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatggcgcaa aacgcctgat ttacgcgag ttccccacag atgatgtgga 300

caagcctggg gataagtgc ctgcggtatt gacattgag gggcgcgact actgacagat 360

gagggcgcg atecttgaca cttagggggc agagtgtga cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgccgt 480

tttccggcga ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaacctg 540

tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccc 600

cttctcgaa cctcccgccc cgctaacyy ggcctcccat cccccaggg gctgogcccc 660

tggccgcga acggcctcac ccaaaaaatg gcagcgctgg cagtccttgc cattgcggg 720

atcggggcag taaogggatg ggcgatcagc ccgagcgca cggccggaag cattgacgtg 780

ccgcagggtc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcgccctg 840

ggtggcgccc tgccttcac ttccggcgtc ggggcattca cggacttcat ggcggggccc 900

gcaattttta ccttgggcat tcttggcata gtggtcgcg gtgccgtgct cgtgttcggg 960

ggtycgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agotaccaag 1080
acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140
agataatata tcttttatat agaagatato gccgtatgta aggatttcag ggggcaaggc 1200
ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260
ctaagtcttg aaaccagga caataacott atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380
agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440
agattccagg tatgcgcctc aattcgctgc gtataatcyl tgcgtattac gtgcagcttt 1500
cccttcaggc gggattcata cagcgccag ccattcgtoa tccatatcac cacgtcaaaag 1560
ggtagacgca ggctcataag acgcccagc gtgcctatag tgcgttcacc gaatacgtgc 1620
gcaacaacgg tcttcggag actgtcatal ccgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttctccatt tcgcgcaga cgatgacgc actgcccgcc 1740
tgtatgcgcg aggttacga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcattcaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgcattgttt 1920
tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca 1980
cccgcagct agctgaacag gagggacagc tgatagacac agaagccact ggagcaoctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat cacocataat tgtggtttca 2100
aaatcggctc cgtcgatact atgttatagc ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaagttt tagaatgcaa ggaacagtga attggagttc gtcttggtat 2220

aattagcttc ttggggatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatqaga atatcacogg aattgaaaaa actgatcgaa aaataccgct cggtaaaaga 2340

tcaggaaaga atgtctcctg ctaaggata taagctggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagy gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagta ggccgatggc gtcctttgct cggagagta 2580

tgaagatgaa caaagccctg aaaagattat cgaagctgat gggagtgca tcaggctctt 2640

tcactccatc gacatatcgg atgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttaactgaata aagatctggc cgaatggat tgcqaaaact gggagaaga 2760

cactccattt aaagatccgc cgaagctgta tgattttta aagacggaaa agcccggaaga 2820

ggaacttgtc ttttccacg ggaacctggg agacagcaac atcttttgta aagatggcaa 2880

agtaagtggc ttatttgatc ttgggagaag cggcagggcg gacaagtgtg atgacattgc 2940

cttctcgctc cgttcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttcgcac caagtgtttt ggtotcagg ccgagggcca cggcaagtat ttgggcaagg 3180

ggtcgtggtt atctgtgcag ggcaagatc ggaataccaa gtacgagaag gacggccaga 3240

cgtctacgg gaccgaattc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gggggtcaaa tcaggataaa gggcacattg ccccggtgtg agtcggggca atcccgcaag 3360

gagggtgaat gaatcggacg tttagccgga aggcatacag gcaagaactg atcgacggg 3420

ggttttcgcg cagggatgcc gaaacctcgc caagcgcac cgtcatgctg ggcggcgcg 3480

aaaccttcca gtcggtcggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540
gcgtgcgaact ggcctccccc geectgcggc cggcatcggc cgcggtggag cgttcgcgtc 3600
gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cagggaacta 3660
tgacgaccaa gaagcgaaaa accgcgggcg aggacctggc aaaacaggtc agcgaggcca 3720
agcaggccgc gttgctgaaa caccagaagc agcagatcaa ggaaatgcag ctttcttgtt 3780
tcgatattgc gccgtggcgc gacacgatgc gagcgatgcc aaacgacacg gccgcctctg 3840
ccctgtttac cagcgcaac aagaaaatcc cgcgcgaggc gctgcaaac aaggtcattt 3900
tcacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggt gtggcgacg gtgttgagtc acgcyaagcg cccccctc ggcgagcgca 4020
tcaccttcac gttctacagc ctttgcaggc acctgggctg gtcgatcaat ggcgggtatt 4080
acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140
accggtttgg gcaacctggaa tcggtgtcgc tgcctgcaccg cttcgcgcgc ctggaccgtg 4200
gcaagaaaaa gtcccgcttc caggctcctga tcgacgagga aatcgctcgt ctgtttgctg 4260
gcgaccacta cagcaaatc atatgggaga agtacggcaa gctgtcggc acggcccgcg 4320
ggatgttoga ctatttcagc tcgcaccggg agcgtgtacc gctcaagctg gaaaccttc 4380
gcctcatgtg cggatcggtt tccaccggc tgaagaagtg gcgcgagcag gtcggcgaa 4440
cctgcgaaga gttgcgagc agcggccttg tggaacacgc ctgggtcaat gatgacctg 4500
tgcaattgcaa acgtagggc cttgtggggt cagttccggc tgggggttca gcagccagc 4560
ctttactggc atttcaggaa caagcgggca ctgctcgacg caettgcttc gctcagtac 4620
gclcggyacy caccggcgcg totacgaact gcgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcagatt tccgcgagat 4740
cggattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcceta 4860
catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttctg ggagcccgaa cagcgaggcc ggggggtcgc 4980
cggtatgtcg ctgcgggctg tgccggcggg tttattgtct gtgatgatcg tccgacagat 5040
tccaaaggga atctggtgga tgcgcattct catctcggc gcacttaata tttcgtctatt 5100
ctggagcttg ttgttttatt cggctctacc cctgcggggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtccctgggg ctatttgcgg aactgcgggc gtggcgctgt tgggtgtgac 5280
accaaacgca ggcctagatc ctgtcggcgt cgcagcgggc ctggcggggc cggtttccat 5340
ggcgttcgga accgtgtcga ccgcgaagt gcaacctccc gtgcctctgc tcacctttac 5400
gccttgcaaa ctggcgccgc gaggacttct gctcgttcca gtacgtttag tgtttgatcc 5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggttttctca gcccagatc tggggctgat cagccgggga tgcacaggc 5640
cgacagtccg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttccat tctaaagaaa tagccgcaat cagcttcctc agcgggttta 5760
tccagcgatt tcttattatg tcggcatagt tctcaagatc gacagcctgt caccgttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctcgc agggagatga tatttgatca 5880
caggcagaa cgctctgtca togttacaat caacatgcta ccctcgcga gatcatccgt 5940

gtttcaaac ccgacagctta gttgccgttc ttccgaatag catcggtaac atgagcaaac 6000
ttgcgcgcct tacaacggct ctccgcgtga cgcgcgtccg gactgatggg ctgcctgtat 6060
cgagtgggtga ttttgtgcg agctgccggg cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaactta taacacattg cggacgtttt 6180
taatgtactg ggggtggttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cgggtccagc tggtttgccc cagcaggcga 6300
aaatcctggt tgatgggtgt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaagggcga aaaaccgtct atcagggcga tggccacta cgtgaacct 6480
caccocaaac aagttttttt gggtcgaggt gccgtaaagc actaaatcgg aacocataag 6540
ggagccccc atttagagct tgacggggaa agcggcgaa cgtggcgaga aaggaaggga 6600
agaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgacgtgtg 6660
cgggcctctt cgctattaag ccagctggcg aaaggggat gtgctgcaag gcgattaagt 6720
tgggtaacgc cagggtttto ccagtcnaga cgttctaaaa cagcggccag tgaattcgag 6780
ctcggtaccc ggggatcttt cgacactgaa atacgtcgag cctgctccg ttggaagcgg 6840
cgaggagcct cgtcctgtca caactaccaa catggagtao gataagggcc agttccgcca 6900
gtcattaag agccagttca tgggcgttg catgatggcc gtcatgcac tgtaactcaa 6960
gtacaccaac gctcttctga tccagtcgat catcgcgtga aggcgccttc gaatctggtt 7020
aagatccag tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgoc 7080
aacagcttcc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgcggag 7140
aagaactgga ggggtgylgt caaggaggag taagctcttt attgaatcg gaggaaggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
 tagcatagct tcatttggat ttgotttcca ggctgagact ctactttgga gcatagaggg 7320
 tcctttggct ttcaatatct tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380
 tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt 7440
 tcgaaatata tcgggatgtc gaaggcttgg ggcacctgog ttggttgaa ttgaacgtg 7500
 gcactattga tcaccgata gctctgcaa gggcgttgca caatgcaagt caaacgttgc 7560
 taacagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620
 tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680
 gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtggg ctgagtgacc 7740
 atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaatg 7800
 cgtatagcta ccagtggaga agtggaacta aatcgacttc agcaacatct cctggataaa 7860
 ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaatc 7920
 tgtccagatc atgggttgacc ggtgcttgga tottctata gaatcatcct tattcgttga 7980
 cctagctgat tctggaatga ccagagggt catgaattga gcctaaaac cgcgcctcc 8040
 accattttga gaaaatgtg acgaactcgt gagctctgta cagtgcacgg tgactctttc 8100
 tggcatgcgg agagacggag ggcgcagag agaaggcgtg agtaataagc cactggccag 8160
 acagctctgg cggctctgag gtgcagtggg tgattattaa tccgggaccg gcgcgcctc 8220
 cgccccgaag tggaaaggct ggtgtgcccc togttgacca agaattctatt gcatcatogg 8280
 agaatatgga gcttcatoga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
 gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400
 atctggtaaa agattcaaga gatagtacct tctccgaagt aggttagagcg agtaccggcc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
 ctgclllgyc cgggtgatga aacgggaaag gcgcgtcagg agctggccag cgggcgagac 8580
 cgggaacaca agctggcagt cgaccatcc ggtgctctgc actcgacctg ctgaggtccc 8640
 tcagtccctg gtaggcagct ttgccccgtc tgcgcgccg gtgtgtcggc ggggttgaca 8700
 aggtcggtgc gtcagtccaa catttgttc catattttcc tgcctccccc accagctgct 8760
 cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
 tatttccctc aagtaagtac ttgctacat ccatactcca tcttcccat coctattcc 8880
 ttggaacctt tcagttcgag ctttcccat tcatcgagc ttgactaaca gctacccgc 8940
 ttgagcagac atcaccatgc ctgaacacac cgcgcgctct gtgcagaagt tctgtatga 9000
 aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctgctgcttt 9060
 cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120
 ctacaaagat cgtttatgttt atcggcactt tgcctcggcc gcgtccccga ttcggaagt 9180
 gcttgacatt ggggaattca gcgagagcct gacctattgc atctccgcc gtgcacaggg 9240
 tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctcgagc cggtcgcgga 9300
 ggocatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360
 accgcaagga atcgggtcaat acactacatg gcgtgalltc atatgcgcga ttgtgatcc 9420
 ccagtgtat cactggcaaa ctgtgatgga cgacaccgto agtgcgctcg tcgcgcaggg 9480
 tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540
 ggatttcggc tccaacaatg tcttgacgga caatggccgc ataacagcgg tcattgactg 9600
 gagcgaggcg atgttcgggg attccaata cgaggtgcgc aacatcttct tctggaggcc 9660
 gtgggtggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatgcgcg cggctocggg cgtatatgct cgcatttggc cttgaccaac tctatcacag 9780
cttggttgac ggcaatttcg atgatgcagc ttggggcgag ggtcgatgcg acgcaatcgt 9840
cogatccgga gccgggactg toggggctac acaaatcgcc cgcagaagcg cggccgtctg 9900
gaccgatggc tgtgtagaag tactgcgca tagtggaaac cgacgcccc gcactcgcc 9960
gaggcgaaag gaatagagta gatgcgcacc gccgggatcg tccacttaac gttactgaaa 10020
tcatacaaca gcttgacgaa tctggatata agatcgttgg tgcgatgctc agctccggag 10080
ttgagacaaa tgggtgtcag gatctcgata agatacgttc atttgtcaa gcagcaaa 10140
gtgccttcta gtgatttaac agctccatgt caacaagaat aaaaacggtt ttccgggtta 10200
cctcttcacg atacagctca tetgcattgc attaatgcat tgactgcaac ctagtcaacg 10260
cttcaggct cgggcgaaga gaagaatagc ttagcagagc tatttctoatt ttccgggagc 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt cgtcaaccag cncctgggtt cgcaaagata attgcatgtt 10500
tcttcttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctcg 10620
taacacccaa tacgucggcc gaaactttt tacaantctc ctatgagtcg ttaccacaga 10680
atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctgcgtgtact 10740
gtgtaagcgc ccaactocaca tctccactcg acctgcagcg atgcaagctt ctttttgcct 10800
tgtaaatttc tggtaactgc caccaagaaa tatgaggata ttctgtatgt tctcgtggt 10860
agccaaaatg atagcacgtg ataaatgacc accaaatagg acggctaatt gtttgggcac 10920
aatgaggctg aacataacc cctattgggt cactatgggg taaaaagta ccaaataga 10980

ataattgtaa tgaacttaaa agcgagggta gcacccaaaa gtaagttaga ttatcacttg 11040
ggatatggag latgtattta gcaaaagttat aaataatagt caacgcaatt atttgccccc 11100
aactccagta acctttcata aaatgaaaat accaagcaaa gaaacttttg tgtttaccat 11160
tgtgaaaatc oggggtctatt gagcttgctg gatttgtgtg gtgtaaccaa tgttttttca 11220
atagtttttg atagtgttaa agaccataaa gggatagggt caatgttcca atcaaatgat 11280
taatcttggt gttttggga aatactacgc catgcatggc atcatgagat gtaataaata 11340
atcccgata taaaatgtt tgccatagta taacaggcaa taacatcaa aattttagct 11400
ttgagatgic aagggaaagt aataaactca ggctaatagc ccatgcgcta acaatgacaa 11460
tagcaatgaa aagccoctta aaclyagatt taottotcag tantggagtc agttttgctt 11520
gatgactgag tggttgttct aactggatca tttctaaga gaaggtggaa caatgttagc 11580
ataattgtgc ttgagttagg actttgaggg taggtacata cttgataaag ttaatgatta 11640
aagagaaaaa aaaagttttg gttcaagca gaaattgttt tttaaatcga ttggtgagaa 11700
aatttttttc tgtttccgca tcaccaaaag cacctcagga atggtcacaa attattggtc 11760
tgattggacc ataagcatac aaaaagtcca ttgaagtata cttagtggct tattagactt 11820
ttatcgtttt ctaacgcgaa tcagcaatgt ttcttgtttg atttactgct tgcttttagat 11880
catttttgtc tgaaatatta tgcatttggt caaagcggcc ttgttttctt ttctttcatg 11940
cttaaacacg ttgtttatc catatattac ttggaatatg catcaccgca aagcgyaagt 12000
gcaaaataac aaagaacctc ttggggttac acgatcaact gctattgtga aaaaaatttc 12060
tttttgaaaa tttttggaat aatatctctt gcaaaaaaga aattttgtat atttagtagc 12120
atcaagaaca aatgaagaa gtgtgggata acaagaatac atcatcttta gacaaaagta 12180

cgagaaaaat ctaataagtt gttatagagg tctttgtttt ctttgtttt atagacagtt 12240
atttagagtt tgaaaagtgt ctctaagtgt tcttttttta ttattattat ttcaaatggt 12300
atgtaataata gctaagacta tagatttgac attttttcta aatataaaat ttacgtcaac 12360
agaaataaat gacacaggtt ctttttctct ctctcaatcc tgttgatcat caatctttga 12420
tgtcgtttta aaacaaatga atggcattta gttcuttagg tgtcaactcac atcttgttga 12480
ccagaaaatc cttattcgcc ctcaaatctg cttttattcct ttcatattgat ttgagtgtta 12540
agtaatgcaa gcaaacaaaa aagaacctt tcttgcaag acaaaagaat tgttttcaga 12600
ggaagcaac tcgttgtcat tttttaagga tttagactta taatcgacac catagtttgt 12660
ccgttacatt ttttattgtc gttttctgat ttctttttaa tctttaagca aaatcaatat 12720
taacttatct tgtcttccaa taaaaaatgg ataccaataa caataaatcc ttccacaaga 12780
aaaaaaaaaa aaactcgaaa aaagcttggc gtaatcatgg tcatagctgt ttctgtgtg 12840
aaattgttat ccgctcacia ttccacacia catacgagcc ggaagcalaa agtgtaaaagc 12900
ctgggggtgc taatgagtga gctaactcac attaatggc ttgcgctcac tgcgcccttt 12960
ccagtcggga aacctgtgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg 13020
cggtttgctg attgggccaa agacaaaagg ggcacattca accgattgag ggagggaaag 13080
taaatattga cggaattat tcattaaagg tgaattatca ccgtcacga cttgagccat 13140
ttgggaatta gagccagcaa aatcaccagt agcaccatta ccattagcaa ggccggaaac 13200
gtcaccaatg aaacatcga tagcagcacc gtaatcagta ggcacagaat caagtttgcc 13260
tttagcgtca gactgtagcg cgttttctac ggcattttcg gtcatagccc ccttattagu 13320
gtttgccatc ttttcataat caaaatcacc ggaaccagag ccaccaccgg aaccgcctcc 13380
ctcagagcgg ccacactcag aaccgccacc ctacagacca ccacctcag agccgccacc 13440

agaaccacca ccagagccgc cgcagcatt gacaggaggc cgcatttagt aacatagatg 13500
acaccgcgcg cgataattta tctagtttg cgcgtatata ttgttttct atgcggtatt 13560
aatgtataa ttgcgggact ctaatcataa aaacccatct cataaataac gtcattgcatt 13620
acatgttaat tattacatgc ttaacgtaat toaacagaaa ttatatgata atcatgcaa 13680
gaccggcaac aggattcaat ctaagaaac ttatttgcca aalytttga cgcgcggga 13740
tcacccgggt ctgtggcggg aactccagca aaatatcga acgcagcaag atatgcgggt 13800
gcattctcgt ctgcctggg cagtcgcgcg cgcgcgcgt gatgtggagc cggggccga 13860
tcattattgc gctcaggatc gtggcgttg gtctgtggc cgttgctgc gtaatgatat 13920
cggcaccttc gaccgcctgt tcgcagaga tccgtgggc gaagaactcc agcatgagat 13980
ccccgcgctg gaggatcac cagccggcgt ccgggaaac gattccgaag ccaaccttt 14040
catagaaggc ggcggtggaa togaatctc gtgatggcag gttggcgtc gcttgctgcg 14100
tcatttcga cccagagtc cgcctcagaa gaactcgtca agaaggcgat agaaggcgt 14160
gcgctgcgaa tcgggagcgc cgataccgta aagcagcagg aagcggtcag cccattgcgc 14220
gcaagctctc tccgcaatat cagcggtagc caacgctatg tccgtatagc ggtccgccac 14280
accagccgcg ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg 14340
caagcagcca tcgccatggc tcacgacgag atcatgcgcg tcgggcgatc gcgccttgag 14400
cctggcgaa agttcggctg gcgcgagccc ctgatctct tcgtccagat catcctgac 14460
gacaagaacc gcttcacatc gagtaoctgc tcgtcgtatg cgaattttgc cttggtggtc 14520
gaatgggcag gtacccgat caagcgtatg cagccgcgcg attgcacag ccatgatgga 14580
tactttctcg gcaggagcaa ggtgagatga caggagatcc tgccccgca cttgcccaa 14640
laguagccag tcccttcgcg cttaagtgc aacgtcagc acagctgcgc aaggaacgc 14700

cgctcgtaggcc agccacgata gccgcgctgc ctctcctgc agttcattca gggcacccga 14760
caggtagctc ttgacaaaa gaaccgggag cccctgcgct gacagccgga acacggcgcc 14820
atcagagcag cagattgtct gttgtgcccc gtcataccg aatagcctct ccaaccaagc 14880
ggcggagaaa cctgcgtgca atccatcttg ttcaatcatg cgaaacgac cagatccggt 14940
gcagattatt tggattgaga gtgaatatga gactctaat ggataccgag ggggaattat 15000
ggaacgtcag tggagcattt ttgacaagaa atatttgcta gctgatagt accttaggag 15060
acttttgac gcgcataat ggtttctgac gtatgtgctt agctcatata actccagaaa 15120
cccgcgctg agtggctcct tcaacgttg ggttctgtca gttccaaacg taaaacggct 15180
tgtcccgct catcgccggg ggtcataacg tgactccctt aattctccg tcatgatcag 15240
attgtcgttt cccgccttca gtttaaaact tcagtgtttg acaggatata ttggcgggta 15300
aacctaagag aaaagagcgt ttattagaat aatcggtat ttaaaaggc gtgaaagggt 15360
ttatccgttc gtccatttgt atgtgcatgc caaccacagg gttccccaga tctggcgccg 15420
gccagcgaga cgagcaagat tggcgccgc ccgaaacgat ccgacagcgc gccacgaca 15480
ggtgngcagc caaattgcac caacgcatac agcgccagca gaatgcata gtggcgggtg 15540
acgtcgttgc agtgaaccag atcgcgagg aggcggcgca gcaccggcat aatcaggccg 15600
atcgccagag cgtcgagcgc gaoagtgtc agaattacga tcagggggat gttgggtttc 15660
acgtctggcc tcgggaccag cctccgctgg tcgattgaa cgcgcggatt ctttatcact 15720
gataagttg tggacatatt atgtttatca gtgataaagt gtaagcatg acaaagttgc 15780
agccaatac agtgatccgt gcgccttg acutgttgaa cgaggtaggc gtacaggttc 15840
tgacgacag caaactggcg gaacgggttg gggttcagca gcggcgctt tactggcact 15900
tcagggaaca gggggcgctg ctgcagcac tggccgaagc catgctggcg gagaatcata 15960

cgcattcgggt gccgagagcc gacgacgact ggcgcctcatt tctgatcggg aatgcccgca 16020
gcttcaggca ggcgcctgctc gcctaccgag atggrgngcg catccatgcc gccacggagc 16080
cgggcgcacc gcagatggaa acggccgacg cgcagcttgg ctctctctgc gaggcgggtt 16140
tttcggcggg ggacgcgcgc aatgcgcgtg tgacaatcag ctacttccact gttggggcgg 16200
tgcttgaggga gcagggccggc gacagcgatg ccggcgagcg cggcggcacc gttgaacagg 16260
ctccgcctctc gccgcctgtg cgggcgcgga tagacgcctt cgacgaagcc ggtccggagc 16320
cagcgttcga gcagggactc gcggtgattg tcgatggatt ggcgaaaagg aggcctcgtt 16380
tcaggaaagt tgaaggaccg agaaagggcg acgattgatc aggaccgctg ccggagcgca 16440
accactcac tacagcagay uatgtagac aacatccctt ccccttctc acngnctag 16500
acgcccgtag cagcccgcta cgggcttttt catgcctctc cctagcgtcc aagcctcacg 16560
gccgcgctcg gcctctctgg cggcctctctg gcgctcttcc gcttctctgc tcaactgactc 16620
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 16680
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 16740
ggccaggaa cgtaaaaagg ccgcgttgct ggcgttttct cataggctcc gccccctga 16800
cgagcatcac aaaaatcgac gctcaagtca gaggtagcga aaccgcagac gactataaag 16860
ataccaggcg tttcccccgt gaagctccct cglcgcctct cctgttccga cctgcgcgt 16920
taccggatac ctgtccgcct ttctcccttc gggaaagcgt gcgcttttcc gctgcataac 16980
cctgcttcgg ggctcattata gcgatttttt ccgtatatcc atccttttcc gcacgataa 17040
caggattttt ccaaaagggt cgtgtagact ttccttggtg tatccaaagg cgtcagccgg 17100
gcaggatagg tgaagtaggc ccaccgcga cggggtgttc cttcttccact gtccttatt 17160
cgcacctggc ggtgctcaac gggaaatcgt ctctgcgagg ctggccgctt accgcggcg 17220

taacagatga gggcaagcgg atggctgatg aaaccaagcc aaccaggaag ggcagccac 17280
ctatcaaggt gtaactgctt ccagcgaac gaagagcgtat tgaggaaaaa ggcgcggcgg 17340
ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca gggctacaaa atcacggcgg 17400
tcgtggacta tgagcacgto cgcgagctgg ccgcacatca tggcgacctg gcccgcctgg 17460
gcggcctgct gaaactctgg ctcaccgacg acccgcgcac ggcgcgggtc ggtgatgcca 17520
cgatcctcgc cctgctggcg aagatogaag agaagcagga cgagcttggc aaggtoatga 17580
tgggcgtggt cgcgccgagg gcagagccat gactttttta gccgctaaaa cggccggggg 17640
gtgcgcgtga ttccaagca cgtcccatg cgctccatca agaagagcga cttcgcggag 17700
ctggtgaagt aatcaccca cgagcaagga aagaacgaga gcctttgcga ngntca 17756

<210> 48

<211> 17118

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n i s a , c , g , o r t

<400> 48

ccgggctgggt tggooctggcc gctgggctgg cggccgtnta tgqccctgca aacgcgccag	60
aaacgcgcgc gaagccgtgt gcgagacacc gggccgcgcg gcgttggtgga tacctgcgcg	120
aaaacttgcc cctcactgac agatgagggg cggacgttga cacttgaggg gcgcactcac	180
ccggcgcggc gttgacagat gaggggcagg ctgcatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat tttaocgcag tttcccacag atgatgtgga	300
caagcctggg gataagtgc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gaggggcgcg atccttgaca cttgaggggc agagtgtctga cagatgaggg gcgcacctat	420
tgacatttga ggggclylcc acgggcagaa aatccagcat ttgcaagggt ttccgcgcgt	480
ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gcgaagggg ggtgcccccc	600
cttctcgaa cctcccgcc cgttaacgcg ggcctcccat cccccaggg gctgcgcgcc	660
tcggccgcga acggcctcac cccaaaaatg gcagcgcctgg cagtccctgc cattgcgggg	720
atcggggcag taacgggatg ggcgacagc ccgagcgca cggccggaag cattgacgtg	780
ccgcaggtgc tggcatcgac attcagcgac cagggtccgg gcagtggagg cggcggcctg	840
ggtggggccc tgccttcac ttgggcgcgc yyygcaltca cggacttcct ggggggggog	900
gcaattttta ccttgggcat tcttggcata gtggtgcgcg gtgcctgtct cgtgttcggg	960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa	1020
acgagaattg gacctttaca gaattactct atgaagccgc atatttaaaa agctaccaag	1080
acgaagagga tgaagaggat gaggaggcag attgccttga atattttgac aatactgata	1140

agataatata tcttttatat agaagatatc gcggtatgta aggatttcag ggggcaaggc 1200
 ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa ctgcatgga 1260
 ctaatgcttg aaaccaggga caataacott atagottgta aattctatca taattgggta 1320
 atgaotcaa ottattgata gtgttttatg ttoagataat gcocgatgac tttgtcatgc 1380
 agctcoaccg attttgagaa cgacagcgac ttccgcccc gccgtgccag gtgotgecte 1440
 agattcagggt tatgcgcgc aattoctgc gtatatogct tgctgattac gtgcagcttt 1500
 cccttcaggc gggattcata cagcgccag ccacccgca tccatatcac cactgcaazg 1560
 ggtgacagca ggctcataag acgccccagc gtgcctatag tgcggtcacc gaatacgtgc 1620
 gcaacaacgg tcttcggag actgtcatc gcgtaaaaca gccagcgctg gcgcgattta 1680
 gccccgacat agccccactg ttgcctcatt tccgcgcaga cgtatgaotc actgcccggc 1740
 tgtatgcgog aggttacoga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800
 ggccaacgcc cataatgcgg gctgttgcgc ggcattccaac gccattcatg gccatatcaa 1860
 tgattttctg gtgcgtacog ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920
 tacqccagtg agagcagaga tagcgtgat gtcggcggt gcttttgcgc ttacgcacca 1980
 ccccgctagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
 aaaaacacca tcatacacta aatcagtaag ttgycagcat caccataat tgtggtttca 2100
 aaatcggtc cgctgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg 2160
 tttctcggtta tttaagttt tagaatgcaa ggaacagtga attggagtc gtctgttat 2220
 aattagcttc ttggggtatc tttaatact gtgaaaaaga ggaaggaaat aataaatggc 2280
 taaaatgaga atatcacogg aattgaaaa actgatogaa aaatacogct gcgtaaaaa 2340
 tacggaaggga atgtctctg ctaaggtata taagctggtg ggagaaaatg aaaaactata 2400

tttaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctcct 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagucga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
ggaacttgtc tttcccacy gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc ttatttgatc ttgggagaag cggcagggcg gacaagtgtg atgacattgc 2940
cttctcgctc cggctgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060
attgttttag tacttagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
tcttcgcgat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggtcgtggtt attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cggctcacg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gcgggtcaaa toaggataaa gggcacattg ccccgcgctg agtcggggca atcccgcaag 3360
gaggggtgaat gaatcggacg ttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc caggatgcc gaaaccatcg caagccgcac cgtcatgctg ggcgcccgcg 3480
aaaccttcca gtcgtcggc tcgatgggtcc agcaagctac ggccaagatc gagcgcgaca 3540
gcgtgcaact ggtcccccct gccctgcccg cgcacatggc cgcggtggag cgttcgcgtc 3600
gtctcgaaac ggaggcggga ggtttgcga atcgaatgac catcgacagc cagggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttctgtaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cagcgccaac aagaaaatcc cgcgcgaggc gctgcaaaac aagggtcattt 3900

tcacagtcac caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgaly 3960

acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtgcgatcaat ggccgggtatt 4080

acacgaaggc cgaggaatgc ctgtcgccgc tacaggcgac ggcgatgggc ttacgtccg 4140

accgcgttgg gcaacctggaa tcggtgtcgc tgcgcaccg cttccgcgc tcggaccgtg 4200

gcaagaaaac gtccgcttgc caggtcctga tcgacgagga aatcgctgtg ctgtttgctg 4260

gcgaccacta cagcaaatc atatgggaga agtacgcga gctgtcgccg acggcccgcac 4320

ggatgttcca ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttc 4380

gcctcatgtg cggatcggtt tccaccgcg tgaagaagtg gcgcgacag gtccggcgaag 4440

cctgcgaaga gttgcgaagc agcggcctgg tggaaacagc ctgggtcaat gatgacctgg 4500

tgcattgcaa acgctagggc cttgtggggt cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cagggcgccg tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccca 4860

catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttctg ggagcccgaa cagcgaggcc gaggggtcgc 4980
cggtatgctg ctacggccgt tgcgggcggg ttatttgcgc gtgatgatcg tccgacagat 5040
tccaaaggga atctgggtga tgcgcatctt catctccggc gcacttaata ttctgcatt 5100
ctggagcttg ttgtttattt cggctatccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgcgcct ctgctaggta gcccgatacg 5220
attgatggcg gtctcggggg ctatttcggg aactcggggc gtggcgctgt tgggtgttac 5280
accaaagcga cgcctagatc ctgtggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga ccgcgaagt gcaacctcc gtgcctctgc tcaactttac 5400
cgctgggcaa ctggcgrrcg gaggacttct cctcgttcca gtacgtttag tgtttgatcc 5460
gcaaaccgg atgcctacag gaaccaatgt totcgccctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc ttgggttcgg ggggatctcg gcactcgaac ctacagtgt 5580
ttccttactg ggcttttcca gccccagatc tggggtcgat cagccgggga tgcatacggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcaact totaaagaaa tagcgccact cagcttcctc agcggcttta 5760
tocagcgatt tctattatg tcggcatagt totcaagatc gacagcctgt cagcgttaag 5820
cgagaaatga ataaagggg tgataattcg gatctctccq agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tegttaaat caacatgcta cctcccgga gatcatcgt 5940
gtttcaaac ccgcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaa 6000
tctgccgct tacaacggct ctcccgctga cgcgctccg gactgatggg ctgcctgtat 6060
cgagtgtga ttttgtgcc agctgccggt cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

laattgtactg ggggtggtttt lcttttcacc agtgagacgg gcaacagctg attgccttc 6240
accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300
aaatcctggt tgatgggtgt tcogaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
cogagatagg gttgagtggt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaaggggcga aaaacogtct atcagggcga tggcccaacta cgtgaaccat 6480
cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aacctaaag 6540
ggagccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaaggga 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgactcgtg 6660
egggcctctt ngctattacg coagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780
ctcggtaacc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840
cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggoc agttccgcc 6900
gtctattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960
gtacaccaac gctctttctg tccagtcgat catcogctga aggcgcttcc gaatctgggt 7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agctccctt taaggctgcc 7080
aacgcgttcc taagccaggc ccagcnaag accgacaaag cctccctcca gaacgccag 7140
aagaactgga ggggtgggtt caaggaggag taagctcctt attgaagtcg gaggacggag 7200
cgggtgcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctgacttgga gcatagaggg 7320
tcctttggct ttcaatatcc tcaagtatct cgagtttgaa cttatccct gtgaaccttt 7380
tattaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatgc catgagggtt 7440

tcgaaataca tccggtatgtc gaaggettggt ggcacctgctg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcggttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggogtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtggaga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactatata gaataagata ggtggagagc ttataccgag ctcccaaatc 7920

tgtccagatc atggttgacc ggtccctgga tcttccatata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga ccagaggggt catgacttga gctctaaaac cgcgcctctc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgcaccg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctcttg cggctctgag gtgcagtggg tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctat gcatactcgg 8280

agaatatgga gcttcatga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagcogt cggcgaata gcatgcactt aactaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagacg agtaccgggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgacagc 8580

cgggaacaca agctggcagt cgaccatcc ggtgctctgc actcgacctg ctgaggtccc 8640

tcagtcocctg gtaggcagct ttgccocgtc tgtcogcccg gtgtgtcggc ggggttgaca 8700
aggctcgttc gtcagtcaca catttggtgc catatcttcc tgctctccccc aocagctgct 8760
ctttctcttt ctctctcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttccoot aagtaagtac ttgtctacat ccaactcca tccttcccat cocttatctc 8880
tttgaacctt tcagttcgag ctttcccatc toatcgagc ttgactaaca gctaccccg 8940
ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtgcagaagt ttctgtatga 9000
aaagttcgac agcgtctccg acctgatgca gctctcggag gggaagaat ctogtgcttt 9060
cagcttcgat gtagggggc gtggatatgt cctcggggta aatagctgag ccgatggttt 9120
ctacaaagat cgttatgttt atcggcactt tgcctccgac gcgctccga ttccggaagt 9180
gcttgacatt ggggaattca gcgagagcct gaactattgc atctccgac gtgcacagg 9240
tgtcacgttg caagacctgc ctgaaccga actgcccgtc gttctgcagc cggctcggga 9300
ggcctatgat gogatcgctc cgcctcatct tagocagac agcgggttcg gccattcgg 9360
accgcaagga atcgttcaat aactacatg gogtgatttc atatgcgca ttgctgatcc 9420
ccatgtgtat cactggcaaa ctgtgatgga cgacacgctc agtgctcgc tcgcgcaggc 9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tctgtcaogc 9540
ggatttcgac tccaacaatg tcctgacgga caatggcgcg ataacagcgg tcattgaotg 9600
gagcaggagg atgttcgggg attccaata cgaggtcgcc aacatcttct tctggaggcc 9660
gtggttggtc tgtatggagc agcagacgag ctacttcgag cggaggcatc cggagcttgc 9720
aggatogcgc cggctccggg cgtatatgct cgcctatggt cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtogatgag acgcaatcgt 9840
cogatccgga gcggggactg tcgggcgtac acaatcgcc cgcagaagcg cggcgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgccca gcactcgtcc 9960
gagggcacaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020
tcatcaaaaca gottgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080
ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtcaa gcagcaaaaga 10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaacgcggtt ttcgggttta 10200
cctcttcag atacagctca tctgcaatgc attaatgat tgaactgaac ctagttaacgc 10260
cttncaggct cgggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgtctctggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctggggt cgcaaagata attgcatggt 10500
tcttctctga actctcaagc ctacaggaca cacattcacc gtaggtataa acctcgaaat 10560
catttctac taagatggtt tacaatagta accatgcacg gttgcctagt gaatgctccg 10620
taacacocaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacagggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctctgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagcgt gagattaaaa 10800
tagataagga aaagaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860
atacatgctt gactttcttt ttccatcgat atgatatatg catatgatag atatacaagc 10920
aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaaagt ttttttttat 10980
acatgtttgt acacaagaat agttaccaat ttgctttggg cttaactgct gcaagtttat 11040
atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100
tttgggagaa ttaggasaag ggagcggaaa ggtaagaaat gcttgcgtat ttactaatt 11160

cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220

tgtattacac tctaatactc gatgacccga ccaaaaagag cgaacaaaga aataatcttg 11280

tgcatctgaa tatgatggaa gattttttcc cctttattct aaatgttgac atagcgtgla 11340

tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt attttatctc 11400

tatgatccag ttagaacaac cactcagtca tcaagcaaaa ctgactccag tactgagaag 11460

taaatctcag ttttaagggc ttttcattgc tattgtcatt gttagcgcgt gggtcattag 11520

cctgagttta ttactttccc ttgacatctc aaagctaaaa ttttggatgt tattgcctgt 11580

tatactatgg caaacatctt tatatacggg attatttatt acatctcatg atgccoatgc 11640

tggcgtagta tttcccaaaa acaccaagat taatcatttg attggaacat tgaccctatc 11700

cctttatggt cttttaccat atcaaaaact attgaaaaaa cattgggtac accaccacaa 11760

tcacgaagc tcaatagacc cggattttca caatggtaaa caccaaagtt totttgcttg 11820

qtatcttcat tttatgaaag gttactggag ttgggggcaa ataattgcgt tgactattat 11880

ttataacttt gctaaatata tactccatat cccaagtat aatctaaact acttttgggt 11940

gctaccctcg cttttaagtt cattacaatt attctatctt ggtacttttt taccccatag 12000

tgaaccaata ggggggtatg ttacagctca ttgtgcocaa acaattagcc gtccatattg 12060

gtggtcattt atcagctgct atcatttttg ctaccaagag gaacatcacg aatatctca 12120

tatttcttggt tggcagttac cagaaattta caagcaaaa tagaagcttg gogtaaatcat 12180

ggcatagct gtttccctgtg tgaattgtt atccgctcac aattccacac aacatacgag 12240

cgggaagcat aaagtgtaaa gcctggggta ctaatgagt gagctaactc acattaattg 12300

cggtgcgctc actgcccgct ttccagtcgg gaaacctgto gtgccagctg cattaatgaa 12360

tcggccaacg cgcggggaga ggcgggttgc gtattggggc aaagacaaaa ggcgcacatt 12420

caaccgattg agggaggga ggtaaatatt gacggaaatt attcattaaa ggtgaattat 12480
caccgtcacc gacttgagcc atttgggaat tagagccagc aaaatcacca gtacaccat 12540
taccattagc aaggccggaa acgtcaccaa tgaaccatc gatagcagca ccgtaatcag 12600
tagcgacaga atcaagtttg cctttagcgt cagactgtag cgcgttttca tcggcatttt 12660
cggctcatagc ccccttatta gcgtttgcca tcttttcata atcaaatca ccggaaccag 12720
agccaccacc ggaaccgect ccctcagagc cgcaccctc agaaccgcca cctcagagc 12780
caccaccctc agagccgcca ccagaaccac caccagagcc gccgcccagca ttgacaggag 12840
gcccgatcta gtaacataga tgacaccgag cgcgataatt tatctagatt tgcgcgctat 12900
atcttgtttt ctatcgctga ttaaatgtat aattgcggga ctctaatcat aaaaacccat 12960
ctcataaata acgtcatgca ttacatgta attattacat gcttaacgtt attcaacaga 13020
aattatatga taatcatgc aagaccgcca acaggattca atettaagaa actttattgc 13080
caaatgtttg aacgatcggt gatcatccgg gtctgtggcg ggaactccac gaaatatcc 13140
gaacgcagca agatatcggt gtgcctctcg gtcttgcctg ggcagtcgac gccgacgccg 13200
ttgatgtgga cgcggggccc gatcatattg tcgctcagga tcgtggcggt gtgcttgtcg 13260
gccgttgcgt tcgtaattgat atcggaacct tcgaccgect gttccgcaga gatcccggtg 13320
gcgaagaact ccagcatgag atccccgcgc tggaggatca tccagccgga gtcccggaaa 13380
acgattccga agcccaacct ttcatagaag gcggcggtgg aatcgaaatc tcgtgatggc 13440
aggttgggag tcgcttgggc ggtcatttcg aacccagag tcccgcctcg aagaactcgt 13500
caagaaggcg atagaaggcg atgcgctgag aatcgggagc ggcgataccg taaagcacga 13560
ggaagcggtc agccattcg ccgccaagct cttcagcaat atcagggta gccaacgcta 13620
tgtctgata gcggtccgca acaccagcc ggcacagtc gatgaatcca gaaagcggc 13680

cattttccac catgatatto ggcaagcagg catgccatg ggtcacgacg agatcatcgc 13740
cgtcgggcat gcgcgccttg agcctggcga acagttcggc tggcgcgagc cctgatgct 13800
cttcgtccag atcatcctga tcgacaagac cggcttccat ccgagtacgt gctcgtcga 13860
tgcgatgttt cgtctgttg tcgaatgggc aggtagccgg atcaagcgta tgcagccgcc 13920
gcattgcacg agccatgatg gatactttct cggcaggagc aaggtgagat gacaggagat 13980
cctgccccgg cacttcgcc aatagcagcc agtcccttcc cgtctcagtg acaacgtcga 14040
gcacagctgc gcaaggaac cccgtcgtgg ccagccacga tagccgcgt gctcgtcct 14100
gcagttcatt cagggcacgg gacaggtcgg tcttgacaaa aagaaccggg cgccctcgcg 14160
ctgacagccg gaacacggcg gcatcagagc agccgattgt ctgttgtgcc cagtcatagc 14220
cgaatagcct ctccacccaa gcggccggag aacctgcgtg caatccatct tgttcaatca 14280
tgcgaaacga tcagatccg gtgcagatta tttggattga gagtgaatat gagactctaa 14340
ttggataccg aggggaattt atggaaacgtc agtggagcat ttttgacaag aaatatattgc 14400
tagctgatag tgaccttagg cgaactttga acgcgcaata atggtttctg acgtatgtgc 14460
ttagctcatt aaactccaga aaccgcggc tgagtggctc cttaacgctt gcggttctlyl 14520
cagttccaaa cgtaaaacgg ctgttccgcg gtcacggcg ggggtcataa cgtgactccc 14580
ttaattctcc gctcatgac agattgtcgt tccccgctt cagtttaaac tatcagtggt 14640
tgacaggata tattggcggg taaacctaa agaaaagagc gtttattaga ataatcggat 14700
atttaaaagg gcgtgaaaag gtttatccgt tcttccattt gtatgtgcat gccaacaca 14760
gggttcccca gatctggcgc cggccagcga gacgagcaag attggccgcc gccggaacg 14820
atccgacagc gcgccacga caggtgcgca ggcaaattgc accaacgcat acagcgccag 14880
cagaatgcca tagtggcgcg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg 14940

cagcaccggc ataatcaggc cgaatccgac agcgtcgagc gcgacagtgc tcagaattac 15000
gatcagggggt atgttgggtt tcacgtctgg cctccggacc agoctccgct ggtccgattg 15060
aacgcgcgga tttcttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa 15120
gtgtcaagca tgacaaagtt gcagccgaat acagtgatcc gtgccgcctt ggaacctgtt 15180
aacagggtcg gcgtagacgg tctgacgaca cgcaaaactg cggaacgggt gggggttcag 15240
cagccggcgc tttactggca cttcagggaac aagcggggcg tgctcgagcg actggccgaa 15300
gccatgctgg cggagaatca lacgattcg gtgccgagag ccgaacgaga ctggcgctca 15360
tttctgatcg ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgaaggcgcg 15420
cgcatccatg ccggcacgcg accgggcgca ccgcagatgg aaacggccga ccgcgagctt 15480
cgcttctctc gcgaggcggg tttttcgcc cgggacgcgc tcaatcgctc gatgacaatc 15540
agctaactta ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgcggcgag 15600
cgccggcgca ccgttgaaca ggctccgctc tcgcgcctgt tgcgggcgcg gatagacgc 15660
ttcgacgaag ccggtcggga cgcagcgttc gagcagggac tcgcggtgat tgtcgatgga 15720
ttggcgaaaa ggaggctcgt tgtcaggaa cgttgaggac cgagaaaggg lyacgattga 15780
tcaggaccgc tgcgggagcg caaccactc actacagcag agccatgtag acaacatccc 15840
ctcccccttt ccaccgcgtc agacgcctgt agcagccgcg tacgggcttt ttcatgcctt 15900
gccttagcgt ccaagcctca cggcgcgcgt.cggcctctctt ggccggcctto tggcgctctt 15960
ccgcttctct gctcaactgac tcgctgcgct cggctgttgc gctgcggcga ggggtatcag 16020
ctcaactcaa ggccgtaata cggttatcca cagaatcagg ggataacgca ggaagaaca 16080
tgtgagcaaa aggcagcaa agggccaggc accgtaaaaa ggccgcgttg ctggcgctttt 16140
tccataggct ccgccccctt gacgagcctc acaaaaatcg acgctcaagt cagaggtggc 16200

gaaacccgac aggactataa agataccagg cgtttccccc tgggaagctcc ctctgtcgct 16260

 ctctctgtcc gacctgcgcg cttacoggat acctgtccgc ctttctccct tcgggaagcg 16320
 tggcgctttt ccgctgcata acctgtctc ggggtcatta tagcgatttt ttccgtatat 16380
 ccactctttt tcgcacgata tacaggattt tgccaaaggg ttctgttaga ctttcccttg 16440
 tgtatccaac ggcgtcagcc gggcaggata gytgaagtag gccaccgcg gagcggtgt 16500
 tccttcttca ctgtccctta ttgcacctg gcggtgtcga acgggaatcc tgctctcgca 16560
 ggctggccgg ctacgcggcg cgtaacagat gagggcaagc ggatggctga tgaaccaag 16620
 ccaaccaggga agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg 16680
 attgaggaaa agggggcggc gcccgcatg agcctgtcgg cctacctgct gcccgctcgc 16740
 cagggtaca aaatcaaggc cgtcgtggac tatgagcagc tcgcgagct gcccgcatc 16800
 aatggcgacc tggccgcctt gggccgcctg ctgaaactct qccctaccga cgaaccgcgc 16860
 acggcgcggt tcggtgatgc cagatcctc gccctgtcgg cgaagatcga agagaagcag 16920
 gacgagcttg gcaaggtcat gatggcgctg gtccgcccga gggcagagcc atgacttttt 16980
 tagccgctaa aacggccggg ggggtgcgct gattgccaa gacgtccca tgcgtccat 17040
 caagaagagc gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga 17100
 gcgcctttgc gacgtca 17118

<210> 49

<211> 18449

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>
<221> misc_feature
<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3679)..(3679)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3770)..(3770)
<223> n is a, c, g, or t

<400> 49
gatcttttoga cactgaaata cgtcgagcct gctcgccttg gaagcggcga ggagcctcgt 60
cctgtcacaac ctaccaacat ggagtaogat aagggcaggt tcggccagct cattaagagc 120
cagttcatgg gcgttggcat gatggcgcgc atgcctctgt acttcaagta caccaacgct 180
cttctgatcc agtogatcat ccgctgaagg cgttttogaa tctggttaag atccacgtct 240
tcgggaagcc agcactgggt gacctccagc gtccttttaa ggcgcgaac agcttttctca 300
ggcagggccca gcccaagacc gacaaggcct cctccagaa cgcgcgagag aactggaggg 360
gtggtgtcaa ggaggagtaa gtccttatt gaagtcggag gacggagcgg tgtcaagagg 420
atattcttcg actctgtatt atagataaga tcatgaggaa ttggaggtag catagcttca 480
tttgatttg ctttcaggc tgagactcta gcttggagca tagagggtcc tttggcttcc 540
aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaaccaatga 600
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc 660
ggatgtcgaa ggccttgggc acctgcgttg gttgaattta gaaogtggca ctattgatca 720

tccgatagct ctgcaaaggg cgttgacaa tgcaagtcaa acgttgctag cagttccagg 780
 tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840
 caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg 900
 caccgctacg gaagacggag aagccacott cagtggactc gagtaccatt taattctatt 960
 tgtgtttgat cgagacctaa tacagccctt acaacgacca tcaaagtcgt atagctacca 1020
 gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa 1080
 ctatacagaa taagataggt gyayayctta taccgagctc ccaaatctgt ccagatontg 1140
 gttgaccggt gcctggatct toctatagaa tcctccttat tcgttgacct agctgattct 1200
 ggagtgacct agagggatcat gacttgagcc taaaatccgc cgcctccacc atttgtagaa 1260
 aaatgtgacg aactcgtgag ctctgtacag tgacgggtga ctctttctgg catgcggaga 1320
 gacggacgga cgcagagaga agggctgagt aataagcoac tggccagaca gctctggcgg 1380
 ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440
 aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcactggaga atatggagct 1500
 tcactgaatc accgacgta agcgaaggag aatgtgaagc caggggtgta tacyuylcgg 1560
 cgaaatagca tgocattaac ctagggtacg aagccaatt gcttcgactc tggtaaaaga 1620
 ttcacgagat agtacctctt cgaagtagg tagagcagat acccgccgag taagctccct 1680
 aattggccca tcggcatctt gtggggcgtc caaatatcgt gcctctcctg ctttgccggg 1740
 tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800
 tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtcctggta 1860
 ggcagctttg ccccgctctg cggcccggtg tgcggcggg gttgacaagg tcgttgctgc 1920
 agtccaacat ttgttgcaat attttcctgc tctccccacc agctgctctt tctttttctc 1980

ttttttttcc catcttcagt atattcatct tcccatccaa gaacctttat ttcccctaag 2040

taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100

gttcgagctt tcccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160

accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gtgcgacagc 2220

gtctccgacc tgatgcagct ctccgagggc gaagatctc gtgctttcag cttcgatgta 2280

ggaggggctg gatatgtcct gccgggtaaat agctgcgcgc atgggtttcta caaagatcgt 2340

tatgtttatc ggcactttgc atcggucygc ctcccgatcc eggaagtgc tgcattggg 2400

gaattcagcg agagcctgac ctattgcac tcccgcctg cacagggtgt cactgttcaa 2460

gacctgcctg aaacgcgaact gccgcctgtt ctgcagccgg tcgcggaggg catggatgcy 2520

atcgctgcgg ccgatcttag ccagacgagc ggggttcggcc cattcgacc gcaaggaato 2580

ggccaataca ctacatggcg tgatttcata tgcgcgattg ctgatccca tgtgtatcac 2640

tggcaaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggtctc cgatgagctg 2700

atgctttggg ccgaggactg ccccgaaagc cggcacctcg tgcacgcgga ttctggctcc 2760

aacaatgtcc tgacggacaa tggccgcata acagcggcca tlyactggag cagggcgatg 2820

ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880

atggagcagc agacgcgcta cttcgagcgg aggcacccgg agcttgcagg atcgccgcgg 2940

ctccggggctg atatgtccg cattggctct gaccaactct atcagagctt ggttgacggc 3000

aatttcgatg atgcagcttg ggcgcagggc cgatgcgacg caatcgccg atccggagcc 3060

gggactgtcg ggcgtacaca aatcgccgcg agaagcgcgg ccgtctggac cgatggctgt 3120

gtagaagtac tcgccgatag tggaaaccga cgcgccagca ctctcccgag gycaaaggaa 3180

tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240

tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tcggagttg agacaaatgg 3300

tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaazagatg ccttctagt 3360

atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420

cagctcatct gcaatgcatt aatgcattga ctgcaacctt gtaacgcctt ncaggctccg 3480

gcgaagagaa gaatagotta gcagagctat ttctatttcc gggagacgag atcaagcaga 3540

tcaacggtcg tcaagagacc taagagactg aggaatccgc tcttggtctc acgcgactat 3600

atatttgtct ctaattgtac ttgacatgc tctctctctt taotctgata gcttgactat 3660

gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tcttgaact 3720

ctcaagceta caggacacac attcctcgta ggtataaacc tcgaaatcan ttctactaa 3780

gatggtatac aatagtaacc atgcatggtt gcctagttaa tgctccgtaa caccoatac 3840

gcggcgcgaa acttttttac aactctccta tgagtcgttt acccagaatg cacaggtaac 3900

cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960

ctccacatct ccactcgacc tgcaggcatg caaagcttga gattaaaata gataaggaaa 4020

agaaagttaa aagaaattcg gaagcatggc acattcttct ttttataaat acatgcctga 4080

ctttcttttt ccctogatat gatatatgca tatgatagat atacaagcaa tcttcttcaa 4140

ggagtttgaa attttgtctc ccaggagcaa aaaaaagttt tttttatata atgtttgtab 4200

acaagaatag ttaccaattt gctttggtct tacgtgtctc aagtttatat cgttttcaat 4260

ttctttgtct ttacatttcc ttgtctctt atctttctct atttagtctt tgggagaatt 4320

aggaaaaggg agcggaaagg taagaatgc ttgcgtattt tactaattcg gcaaacatcc 4380

aatttggcaa acagcagcct gtgcaacgct ctcgagatga cagtatcttt gattacactc 4440

taaatctcga tgaccgcgacc aaaaagagcg aacaaagaaa taatcttggt cattcgaata 4500

tgatggaaga ttttttcccc cttattctaa atgttgacat agcgtgtatg ttatataaac 4560

aaaaagaaat tgtacaaact ttcttttctt ctctttttat ttatctctta tgctgtogaa 4620

gctgcagtc aatcagcgtca aggcccgccg cgttgaaact gcccgcgaca tcacgcggcc 4680

caaaactctgc ctgcatgctc agcgggtgctc gttagtctgg ctgcgagtgg cagcaccaca 4740

gacagaggag gcgctgggaa ccgtgcaggg tcgcggcgcg ggcgatgagc acagcgccga 4800

tgtagcactc cagcagcttg accgggctat cgcagagcgt cgtgcccggc gcaaaccgga 4860

gcagctgtca taccaggctg ccgccattgc agcatcaatt ggcgtgtcag gcattgccat 4920

cttcgccacc tacctgagat ttgccatgca catgacogtg ggcggcgagc tgccatgggg 4980

tgaagtggct ggcactctcc tcttggtggt tgggtggcgcg ctcggcatgg agatgtatgc 5040

ccgctatgca cacaaagcca tctggcatga gtgcctctg ggcgtgctgc tgcacaagag 5100

ccaccacaca cttgcactg racctttga agccaacgac ttgtttgcaa tcatcaatgg 5160

actgcccgc atgctcctgt gtacctttgg cttctggctg cccaacgtcc tggggggcgc 5220

ctgcttttga ggggggtggt gcatacagct atacggcatg gcataatatgt ttgtacacga 5280

tggcctgggt cacaggcgct ttcccacogg gccatcgct ggcctgccct acatgaagcg 5340

cctgacagtg gccaccagc tacaccacag cggcaagtac ggtggcgcg cctggggtat 5400

gttcttgggt ccacaggagc tgcagcacat tccaggtgcg gcggaggagg tggagcgact 5460

ggtcctggaa ctggactggt ccaagcgggc gattgtgact gatagcgaga ctctgggtcg 5520

atgttatctg cctoaacaat ggcttagaaa agaagaazaca gaacaaatac agcaaggcaa 5580

cgcccgtagc ctagggtgac aaagactgtt gggcttgtct ctgaagcttg taggaaggcc 5640

agacgctatc atggtgagag ctaagaaggg cattgacaag ttgccggcaa actgtcaagg 5700

cgggtgtacga gctgcttgcc aagtatatgc tgcaattgga tctgtactca agcagcagaa 5760
gacaacatat cctacaagag ctcactctaaa aggaagcgaa cgtgccaaaga ttgctctgtt 5820
gagtggtatc aacctctatc aatctgaaga caagcctgtg gctctocgtc aagctagaaa 5880
gattaagagt tttttgttg attagtgaat tttgtttta ttatgtctg atagttcaat 5940
aaagaacaa cacatacaat ataaaatcat tgtctttaa tgttaattta gttagtgta 6000
aagcctgcat ttttttgtg cgcataaaca atgaattcac ccgcttctg gtttttaaat 6060
aattatgtca aactaggga aattctttt ttctctctg ttctttttt ggcttggtt 6120
ggagtcaag gctgtcttc agattgatag aggttgata cactcaacag agcaatctg 6180
gcagcttcg ttctttttg atgagctct gttaggatag ttgtctctg ctgcttgagt 6240
acagatcaa ttgcagcata tacttgga gcagctcgta caccgcttg acagtttgcc 6300
ggcaactgt caatgcctt cttagctct accatgatag cgtctgctt tctacaagc 6360
ttcagagaca agccaacag tctttgatc ctaggctac gggcgttgcc ttgctgtatt 6420
tgttctgtt ctcttttct aagccattgt tgaggcagat aacatcgacc caacatctc 6480
gagccatact acagcataaa aggatacgt ttctttaaca gaaatttacc ottttgttat 6540
cagcacatc aaaaaaaay aaatttaaga tgagtaggac ttccattctc tcaaaaattt 6600
tattcaatc ataatgaat ttttttga caaaaagaa agattatgoc tgattttctc 6660
tattttttt ttttttcaa ctcaaccaat actttctagc ccagcttgcc gtaatcatg 6720
tcgatagct ttctgtgtg aaattgttat ccgctcaca ttccacacaa catacgagcc 6780
ggaagcata agtgtaaagc ctgggggtgc taatgagtga gctaactcac attaatgoc 6840
ttgcgtcac tgcctgctt ccagtcggga aactgtcgt gccagctgca ttaatgaatc 6900
ggcaacgcg cggggagagg cggtttgcgt attgggcaa agacaaaagg gcacattca 6960

accgattgag ggaggggaagg taaatattga cggaaattat tcattaaagg tgaattatca 7020
 ccgtcacoga cttgagocatt ttgggaatta gagccagcaa aatcaccagt agcaccatta 7080
 ccattagcaa ggccggaaac gtcaccaatg aaacocatga tagcagcacc gtaatcagta 7140
 ggcacagaat caagttttgcc tttagcgtca gactgtagcg cgttttcacg ggcattttcg 7200
 gtcataqccc cottattagc gtttgccatc ttttcataat caaatcacc ggaaccagag 7260
 ccaccacogg aaccgcctcc ctacagacgg ccaccctcag aaccgccacc ctacagagcca 7320
 ccaccctcag agccgcacuu ayaaccacca ccagagcccg cgcacgcatt gacaggaggo 7380
 ccgatctagt aacatagatg acacgcgcg cgataattta tcctagtttg cgcgctatat 7440
 tttgttttct atcgcgtatt aaatgtataa ttgcgggact ctaatcataa aaacccatct 7500
 cataaataac gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa 7560
 ttatatgata atcatcgcaa gacgggcaac aggattcaat cttaagaaac ttatttgcca 7620
 aatgtttgaa cgcacgggga tcatccgggt ctgtggcggg aactccacga aaatatccga 7680
 acgcagcaag atatgcgggt gcatctcggt cttgcctggg cagtcgcccg cgacgccgtt 7740
 gatgtggagc cggggccoga tcatattgtc gtcaggatc gtyyuytltt gcttgtcggc 7800
 cgttgtgtgc gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccgtaggc 7860
 gaagaactcc agcatgagat cccgcgctg gaggatcatc cagccggcgt cccggaaaac 7920
 gattccgaag cccaaccttt catagaaggc ggcggtggaa tcgaaatctc gtgatggcag 7980
 gttgggcgtc gcttggtcgg tcatttogaa ccccgagtc ccgctcagaa gaactcgtca 8040
 agaaggcgat agaaggcgat gcgctcgcaa tcgggagcgg cgataccgta aagcacgagg 8100
 aagcggtcag ccattcgcc gccaaactct tcagcaatat caggggtagc caacgtatg 8160
 tcttgatagc ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca 8220

ttttccacca tgatattcgg caagcaggca tggccatggg tcacgacgag atcatcgccg 8280
 tggggcatcg gcgccttgag cctggcgaac agttcggctg gcgcgagccc ctgatgctct 8340
 tggtcagat catcctgac gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg 8400
 cgatgtttcg cttgggtggc gaatgggcag gtagccggat caagcgtatg cagcccgccg 8460
 attgcatcag ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc 8520
 tgccccggca cttgcgcca tagcagccag tcccttcccg cttcagtgac aacgtcgagc 8580
 acagctgcgc aaggaacgcu cgtcgtggcc agccacgata gcgcgctgc ctgctcctgc 8640
 agttcattca gggcaccgga caggtcggc ttgacaaaa gaaccgggcg cccctgcgct 8700
 gacagccgga acacggcggc atcagagcag ccgattgtct gttgtgccca gtcatagccg 8760
 aatagcctct ccacccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg 8820
 cgaaacgac cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt 8880
 ggataccgag gggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta 8940
 gctgatagtg accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt 9000
 agtccattaa actccagaaa cccgcggctg alygctcct tcaacgttgc ggttotgtca 9060
 gtccaaaacg taaaacggct tgtcccgcgt catcgccggg ggtcataacg tgactccctt 9120
 aattctccgc tcctgatcag attgtcgttt ccgccttca gtttaacta .tcagtgtttg 9180
 acaggatata ttggcgggta aacctagag aaaagagcgt ttattagaat aatcggatat 9240
 ttaaaagggc gtgaaaaggt ttatccgttc gtccatttgt atgtgcacgc caaccacagg 9300
 gttccccaga tctggcgccg gccagcgaga cgagcaagat tggcgcgcgc ccgaaacgat 9360
 ccgacagcgc gccacgaca ggtgcgcagg caaattgcac caacgcatac agcgccagca 9420
 gaatgccata gtggcggtg acgtcgttcg agtgaaccag atcgccgagc aggcccgcca 9480

gcaccggcat aatcaggcgc atgccgacag cgtcgagcgc gacagtgtctc agaattacga 9540
 tcaggggtat gttgggtttc acgtctggcc tccggaccag cctccgtctg tccgattgaa 9600
 cgcgcggatt ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt 9660
 gtcaagcatg acaaagttgc agccgaatac agtgcacgtg gcgcgccttg acctgttgaa 9720
 cggggtcggc gtgacggctc tgacgacacg caaactggcg gaacggtttg gggttcagca 9780
 gcggcgctt tactggcact tcaggaaaca gggggcgtg ctcgacgcac tggccgaagc 9840
 catgctggcg gaaalcata cgcattcggc gcgcgagcgc gacgaagaot ggggtcatt 9900
 tctgatcggg aatgcccgca gcttcaggca ggcgtgtctc gcctaccgag atggcgcgcg 9960
 catccatgcc ggcacgcgac cgggcgcacc gcagatggaa acggccgacg cgcagcttcg 10020
 ctctctctgc gaggcgggtt tttcgccggg ggacgcccgc aatgcgttga tgacaatcag 10080
 ctacttccact gttggggccg tgccttgagga gcaggccggc gacagcgatg ccggcgagcg 10140
 cgggcgcacc gttgaacagg ctccgctctc gcgcgtgttg cgggcgcgca tagacgcctt 10200
 gcacgaagcc ggtccggacg cagcgttcga gcagggaactc gcgggtgattg togatggatt 10260
 ggcgaaaagg aggcctcgtg tcaggaaactc tgaaygacg agaaagggtg acgattgato 10320
 aggaöcgtg ccggagcgca acccactcac tacagcagag ccattgtagac aacatccctt 10380
 cccctttcc acgcgtcag acgcgcgtag cagcccgcta cgggcttttt catgccctgc 10440
 cctagcgtcc aagcctcag gccgcgtcg gcctctcttg cggcctcttg gcgctcttcc 10500
 gcttctctgc tcaactgact gctgcgctcg gctgttcggc tgcggcgagc ggtatcagct 10560
 cactcaaaag cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 10620
 tgagcaaaag gccagcaaaa ggccagggaac cgtaaaaagg ccgcgttgtct ggcgtttttc 10680
 cataggctcc gcccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 10740

aaccgcagag gactataaag ataccaggcg ttccccctg gaagctccct cgtgcgctct 10800

cctgttcoga cctgcgcgt taccggatac ctgtccgct ttctcccttc gggaagcgtg 10860

gcgcttttcc gctgcataac cctgcttcgg ggtcattata gcgatttttt cggatatatcc 10920

atcctttttc gcaagatata caggattttt ccaagggtt cgtgtagact ttccttggtg 10980

tatccaacgg cgtcagccgg gcaaggatagg tgaagtaggc ccaccgcga gccgggtgttc 11040

cttcttcaact gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg 11100

ctggccgggt accyccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc 11160

aaccaggaag ggcagcccac ctatcaagggt gtactgcctt ccagacgaac gaagagcgat 11220

tgaggaaaag gccgcggcgg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca 11280

gggctacaaa atcacggggc tctgggacta tgagcacgtc ccgcagctgg ccgcatcaa 11340

tgggcacctg gccgcctcgg ccgcctcgt gaaactctgg ctacccgaag acccgcgcac 11400

ggcgcgggtc ggtgatgcca cgatcctcgc cctgtcggcg aagatcgaag agaagcagga 11460

cgagcttgcc aaggtcatga tgggcgtggt ccgcgcgagg gcagagccat gactttttta 11520

gcgcgtaaaa ccgcgggggg gtgcgcgtga ttgccaagca cgtcccatg cgtccatca 11580

agaagagcga cttcgcggag ctggtgaagt acatcacga cgagcaaggc aagaccgagc 11640

gccttttgca cgtcacccgg cctggttgcc ctgcgcctg ggtggcggc cgtctatggc 11700

cctgcaaacg cgcagaaac gccgtcgaag ccgtgtcga gacaccgcgg ccgcggcgt 11760

tgtggataco tccggaaaaa cttggccctc atgacagat gaggggcga cattacact 11820

tgagggggcc actcaccggc cggggcgtt acagatgagg ggcaggctcg atttcggccg 11880

gcgacgtgga gctggccagc ctgcgaatc ggcgaaaac cctgatttta cgcgagtttc 11940

ccacagatga tgtggacaag cctggggata agtgcctgc ggtattgaca cttagagggc 12000
gcgactactg acagatgagg ggcggatcc ttgaccttg aggggcagag tgctgacaga 12060
tgagggggcg acctattgac atttgagggg ctgtccacag gcagaaaac cagcatttc 12120
aagggtttcc gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca 12180
atatttataa accttgtttt taaccagggc tgcgcctgt gcgctgacc gcgcacgcgc 12240
aaggggggtg ccccccttc tcgaacctc ccggccgct aacgcgggccc tcccatcccc 12300
ccaggggctg cgcctctcg ccgcgaacgc cctcaccca aaaatggcag cgctggcagt 12360
ccttgccatt gcgggatcg gggcagtaac gggatgggc atcagccga gcgcgacgc 12420
cggaagcatt gacgtgcgc aggtgctggc atcgacattc agcgaccagg tgccgggcag 12480
tgaggggggc ggcctgggtg gcggcctgcc cttaacttc gccgtcgggg cattoacgga 12540
cttcattgcg ggcgcggcaa tttttacctt gggcattctt ggcatagtgg tcgcgggtgc 12600
cgctcctgtg ttcgggggtg ccataaaccc agcgaacctt ttgaggtgat aggtaagatt 12660
ataccgaggt atgaaaacga gaattggacc ttacagaat tactctatga agcgccatat 12720
ttaaaaagct accaagacga agaggatgaa gaggatgagg aggcagattg ccttgaatat 12780
attgacaata ctgataagat aatatatctt ttatataga gatatgccg tatgtaagga 12840
tttcaggggg caaggcatag gcagcgcgt tatcaatata tctatagaat gggcaaacga 12900
taaaaaactg catggactaa tgcttgaac ocaggacaat aacottatag cttgtaaatt 12960
ctatcataat tgggtaatga ctccaacta ttgatagtgt tttatgttca gataatgcc 13020
gatgaacttg toatgcagct ccacgattt tgagaacgac agcgacttcc gtcccagccg 13080
tgccaggtgc tgctcagat tcaggttatg ccgtcaatt cgctgcgtat atcgctgtct 13140
gattacgtgc agctttccct tcaggcggga ttcatacgc gccagccat ccgtcatcca 13200

tatcacocag tcaaaggggtg acagcaggct cataagacgc cccagcgctg ccatagtgcg 13260

ttcacccaat acgtgcgcaa caacogtctt cggagagctg tcatacgcgt aaaacagcca 13320

gcgctggcgc gatttagccc cgacatagcc coactgttcg tccatttcgc cgcagcagat 13380

gagctcactg cccggctgta tgcgcgaggt taccgactgc ggctgagtt ttttaagtga 13440

cgtaaaatcg tgttgaggcc aacgcccata atgcgggctg ttgcccggca tocaacgcca 13500

ttcatggcca tatcaatgat tttctggcgc gtaccgggtt gagaagcggg gtaagtgaac 13560

tgcagttgcc atgttttaag gcagtgagag cagagatagc gctgatgtcc ggcggtgctt 13620

ttgccgttac gcaccacccc gtcagtagct gaacaggagg gcagctgat agacacagaa 13680

ggcactggag caccctaaaa acaccatcat acactaaac agtaagttgg cagcatcacc 13740

cataattgtg gtttcaaaaat cggctccgct gatactatgt tatacgccaa ctttgaanaa 13800

aactttgaaa aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg 13860

gagttcgtct tgttataatt agctttctgg ggtatcttta aatactctag aaaagaggaa 13920

ggaaataata aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat 13980

accgctgctg aaaaagatag gaaggaatgt ctctgctaa ggtatataag ctggtgggag 14040

aaaaagaaaa cctatatatta aaaatgacgg acagccggta taaagggacc acctatgat 14100

tggaacggga aaaggacatg atgctatggc tggaaggaaa gctgctgtt ccaagggtcc 14160

tgcactttga accgcatgat ggctggagca atctgctcat gagtgaggcc gatggcgctc 14220

tttgctcgga agagtatgaa gatgaacaaa gcctgaaaa gattatcgag ctgtatcgcg 14280

agtgcatcay gctctttcac tccatcgaca tatcgagttg tccctatcag aatagcttag 14340

acagccgctt agcogaattg gattacttac tgaataacga totggccgat gtggattgcg 14400

aaaactggga agaagacact ccatitaaag atccgcgcga gctgtatgat tttttaaga 14460

cggaaaagcc cgaagaggaa ctgtctttt cccacggcga cctgggagac agcaacatct 14520
ttgtgaaaga tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcgggaca 14580
agtggatga cattgccttc tgcgtccggt cgcacggga ggatcggg gaagaacagt 14640
atgtcgagct attttttgac ttactgggga tcaagcctga ttgggagaaa ataaatatt 14700
atattttact ggaatgaattg ttttagtacc taatgttggc gcaacgatgc cggcgacaag 14760
caggagcgca cgcactcttc ccgcacaaag tgttttggct ctccggcgga gggccacggc 14820
aagtatttgg gcaaggggtc gctggtattc gtgcagggca agattcggaa taccaagta 14880
gagaaggacg gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg 14940
gacaccaagg caccaggcgg gtcaaatcag gaataagggc acattgcccc ggcgtgagtc 15000

ggggcaatcc cgaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa 15060
gaaactgatcg acggggggtt ttccgcccag gatgccgaaa ccacgcgaag ccgcaccgtc 15120
atgcgtgcgc cccgcgaaac cttccagtcg gtccgctcga ttgtccagca agctacggcc 15180
aagatcgagc gcgacagcgt gcaactggct cccctgccc tgcccgcgcc atcgccgcc 15240
gtggagcggt cgcgtctctc cgaacaggag gcggcaggtt tggcgaagtc gatgaccatc 15300
gacacgcgag gaactatgac gaccaagaag cgaaaaaacc cggcgaggga cctggcaaaa 15360
caggtcagcg aggccaaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa 15420
atgcagcttt ccttgttcga tattgcgccg tggcgggaca cgatgcgagc gatgcaaac 15480
gacacggccc gctctgccct gttaccacg cgaacaaga aaatcccgcg cggggcgctg 15540
caaaaacagg tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgtcgag 15600
ctcggggccg acgatgagca actggtgtgg cagcaggtgt tggagtacgc gaagcgacc 15660
cctatcggcg agccgatcac cttaacgttc tacgagcttt gccaggacct gggctggctg 15720

atcaatggcc ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg 15780

atgggcttca cgtccgacgc cgttgggcac ctggaatcgg tgcgcctgct gcaccgcttc 15840

cgcgtcctgg accgtggcaa gaaaacgtcc cgttgcacgg tcctgatcga cggagaaatc 15900

gtcgtgctgt ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg 15960

tgcgcgacgg ccgcacggat gttcgaatat ttcagctcgc accgggagcc gtacccgctc 16020

aagctggaaa ccttcgcct catgtgcgga tcggattcca ccgcggtgaa gaagtggcgc 16080

gagcaggctcg gogaagccty cyaaagcttg cgcggcagcg gcctgggtgga acacgcctgg 16140

gtcaatgatg acctgggtgca ttgcaaacgc tagggccttg tggggctcagt tccgctgggg 16200

ggttcagcag ccagcgcctt actggcattt caggaacaag cgggcactgc tcgacgcact 16260

tgcttcgctc agtatcgctc gggacgcacg gcgcgcctca cgaactgcgc ataaacagag 16320

gattaaaatt gacaattgtg attaaggctc agattcgacg gcttggagcg gcgcacgtgc 16380

aggatttcgg cgagatcga ttgtggccc tgaagaaagc tccagagatg ttccgggtccg 16440

tttacgagca cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg 16500

tggcattcgg cgcctacatc gacggcgaga tcattgggct gtcggctctc aaacaggagg 16560

acggcccca ggcgcctcac aaggcgcac tgtccggcgt ttctgtggag ccgaacacg 16620

gaggccgagg ggtcgcggct atgctgctgc gggcgttcc gcggggttta ttgctcgtga 16680

tgatcgtccg acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgac 16740

ttaatatctt gctattctgg agcttgttgt ttatttcggt ctaccgcctg cggggcgggg 16800

tgcggcgac ggtaggcgt gtgcagcgc tgatggtcgt gttcatctct gcgcctcgc 16860

taggtagccc gatacgattg atggcggctc tggggctat ttgcggaact gcggcgctgg 16920

cgtgttgggt gttgacacca aacgcagcgc tagatcctgt cggcgtcgca gcgggcctgg 16980

cgggggcggt ttccatggcg ttcggaaccg tgcagaccg caagtggcaa cctcccgtc 17040

ctctgtctac ctttaccgcc tggcaactgg cggccggagg acttctgtct gtccagtag 17100

ctttagtgtt tgatccgcca atcccgatgc ctacaggaac caatgtttct ggccctggcg 17160

ggctcgccct gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac 17220

togaacctac agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc 17280

cggggatgca tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcggtgag 17340

caalggalag yggagttgat atcgtaacg ttcactttct aagaaataga gcaactcagc 17400

ttctcagcg gotttatoca gcgatttctt attatgtcgg catagtctct aagatcgaca 17460

gcctgtcacg gttaagcgag aatgaataa gaaggctgat aattcggatc tctgcggagg 17520

agatgatatt tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaacct 17580

cgcgagatc atccgtgttt caaacccggc agcttagttg ccgtttcttc gaatagatc 17640

ggtaacatga gcaagctct cgcctttaca accgctctcc cgtgacgcc gtccggact 17700

gatgggctgc ctgtatcgag tgggtatttt gtgcgagct gccggtcggg gagctgttgg 17760

ctggctggtg gcaggatata ttgtgggtga aacaaattga cgtttagaca acttaataac 17820

acattcgga cgtttttaat gtactggggt ggtttttctt ttcaccagt agacgggcaa 17880

cagctgattg ccttcaccg cctggccctg agagagttgc agcaagcgtt ccacgtggt 17940

ttgcccagc aggcgaaat cctgtttgat ggtggttccg aaatcggcaa aatccctat 18000

aaatcaaaag aatagccga gatagggtt agtgtgttc cagtttgaa caagagtcca 18060

ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 18120

ccactacgtg aaccatcac caaatcaagt tttttgggt cgaggtccg taaagcacta 18180

aatcggaacc ctaaaggagg ccccgattt agagcttgac ggggaaagcc ggcgaacgtg 18240

gcgagaaagg aagggaagaa agcgaaagga gcggggcgcca ttcaggctgc gcaactgttg 18300
ggaagggcga tcgggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc 18360
tgcaaggcga ttaagttggg taacgccagg gttttccag tcacgacgtt gtaaaacgac 18420
ggccagtga ttcagctcg gtaccoggg 18449

<210> 50

<211> 18617

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 50

cgggctggt tgcctcgcc gctgggctgg cggcgtcta tggcctgca aacgcgccg 60

aaacgcgctc gaagcgtgt gcgagacacc gcggccgccg gcgttggtga tacctcgagg 120

aaaacttgcc ctcactgac agatgagggg cggacgttga caottgaggg gccgactcac 180

cggcgccggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcgcgaa aacgcctgat ttacgcgag ttcccacag atgatgtgga 300
 caagcctggg gataagtgc ctgcggtatt gacacttgag ggcgcgact actgacagat 360
 gagggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacotat 420
 tgacatttga gggcgtgtcc acaggcagaa aatccagcat ttgcaagggt ttccgccgt 480
 ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg 540
 tttttaacca gggcgtgcgc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcocccc 600
 cllctcgaa cctcccgccc cgttaacggg ggocctccat ccccccaggg gntgngcccc 660
 tggcgcgca acggcctcac ccaaaaaatg gcagcgctgg cagtccctgc cattgcgggg 720
 atcggggcag taacgggatg ggcgatcagc ccgagcgca gcccggaag cattgacgtg 780
 ccgcaggtgc tggcatcgac attcagcgac caggtgccg gcagtgaggg cggcgccctg 840
 ggtggcgccc tgcccttcac ttcgccctgc ggggcattca cggacttcat ggcggggccg 900
 gcaattttta ccttgggcat tcttggcata gtggtgcgg gtgccgtgct cgtgttcggg 960
 ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020
 acgagaattg gacctttaca gaallactct atgaagcgc atatttaaaa agctaocaa 1080
 acgaagagga tgaagaggat gaggaggcag attgccttga atatttgac aatactgata 1140
 agataatata tcttttatat agaagatac gccgtatgta aggatttcag ggggcaaggg 1200
 ataggcagcg cgtttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260
 ctaatgcttg aaaccagga caataacctt atagcttgta aattotatca taattgggta 1320
 atgactccaa cttattgata gtgttttatg ttacagataat gccgatgac ttgtcatgc 1380
 agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440
 agattcaggt tatgcgcctc aattogctgc gtatatcgct tctgallac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatocgtca tccatatcac caggtcaaaag 1560
gggtcacagca gggtcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620
gcaacaacccg tottcoggag actgtcatat gcgtaaaaca gccagcgtcg gcgcgattta 1680
gccccgacat agccccactg ttctgtccatt tcgcgcgaga cgatgacgtc actgcccggc 1740
tgtatgcgcg aggttaccca ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatcggg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttyayaag cytgtaagt gaactgcagt tgcctgtttt 1920
tacggcagtg agagcagaga tagcgtgat gtccggcggg gcttttgcg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcataccta aatcagtaag ttggcagcat cacccataat tgtggtttca 2100
aaatcggctc cgtgcatact atgttatacg ccaactttga aaacaacttt gaaaagctg 2160
ttttctggta ttttaaggtt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220
aattagcttc ttggggatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcacgg aattgaaaaa actgatcgaa aaataccgct gctlaaaaga 2340
tacggaagga atgtctcctg ctaaggata taagctggg ggagaaaatg aaacotata 2400
tttaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgact ttgaacggca 2520
tgatggctgg agcaatctgc tcattgagtga ggccgatggc gtcccttgcg cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtga tcaggctctt 2640
tcactccate gacatatacg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attgattac ttactgaata acgatctggc cgatgtggat tgcgaaaact ggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agccccgaaga 2820
 ggaacttgtc ttttccacg gcgacctggg agacagcaac atcttttgta aagatggcaa 2880
 agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940
 cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000
 tgacttactg gggatcaago ctgattggga gaaaataaaa tatttatattt tactggatga 3060
 attgttttag tacctagatg tggcgcaacg atgcggcgga caagcaggag cgcaccgact 3120

 tcttcgcgat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
 ggctcgtggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
 cgggtctacg gaccgacttc attgcgata aggtggatta totggacacc aaggcaccag 3300
 gcgggtcaaa tcaggaataa gggcacattg ccccgsgctg agtcggggca atoccgcaag 3360
 gagggtgaa: gaatcggacg tttacccga aggcatacag gcaagaactg atocagcggg 3420
 ggttttcgcg cgaggatgcc gaaaccatcg caagccgcac cgtcatgctg gcgccccgcg 3480
 aaaccttcca gtccgtggcg togatgggtc gcaagctac ggccaagato gagcgcgaca 3540
 gogtgcaact ggctccccct gccctgcccc cgcctatggc gcgcgtggag cgtttcgctc 3600
 gtctcgaaca ggagcgggca ggtttgcgga agtcgatgac catcgacaag cgaggaacta 3660
 tgacgaccaa gaagcgaaaa accgcggcgg aggacctggc aaaacaggto agcgaggcca 3720
 agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaatgcag ctttccttgt 3780
 tcgatattgc gcgtggcgg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840
 cctgtttcac caccgcgaac aagaaaatcc cgcgcgaggg gctgcaaaac aaggctattt 3900
 tccacgtcaa caaggacgtg aagatcacct acacggcgct cgagctgggg gccgacgatg 3960

acgaactggt gtggcagcag gtgttgagat acgcgaagcg caccocctate ggcgagccga 4020

tcacettcac gttctacgag ctttgccagg acctgggctg gtogatcaat ggcgggtatt 4080

acacgaaggg cgaggaatgc ctgtgcgcgc tacaggcgac ggcatggggc ttacagtccg 4140

acgcggttgg gaacctggaa tcggtgtcgc tgcgcacgg cttccgcgtc ctggacogtg 4200

gcaagaaaaa gtcccggttc caqutccga tcgacgagga aatcgctgtg ctgtttgctg 4260

gcgaccacta cagcaaattc atatggggaga agtaccgcaa gctgtgcggc acggcccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttc 4380

gcctcatgtg cggatcggat tccaccgcgc tgaagaagtg gcgcgagcag gtccggcgaag 4440

ctcgcgaaga gttgcgaggc agcggcctgg tggaaacagc ctgggtcaat gatgaocctg 4500

tgcattgcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtac 4620

gctegggagc caaggcgccg tctacgaact gccgataaac aqaaqattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtgc gccctgaaga aagctccaga gatgttcggg tccgtttacg agcagcagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcgggccta 4860

catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggcgt tgcggcgggc tttattgtc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcatttc catctcggc gcacttaata tttcgctatt 5100

ctggagcttg ttgtttattt cgggtctaccg cctgcggggc ggggtgcggc cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcac ctctgcggct ctgctaggta gcccgatacg 5220

attgatggcg gtcctggggg ctatttgcgg aactgcgggc gtggcgctgt tgggtgtgac 5280
 accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
 ggcgctcgga accgtgctga ccgcgaagt gcaacctccc gtgcctctgc tcacctttac 5400
 cgctctggca ctggcggcgc gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
 gcaaatcccc atgcctacag gaaccaatgt totcggcctg gcgtggctcg gcctgatcgg 5520
 agcgggttta aactacttcc tttggttccg ggggatctcg cgaactgaac ctacagtgtt 5580
 ttcttactg ggctttctca gcccagatc tggggtgat cagcggggga tgcctcaggc 5640
 cgacagtcgg aacttcgggt .ccccgacctg taccattcgg ttagcaatgg ataggggagt 5700
 tgatatcgtc aacgttctact tctaaagaaa tagcgcact cagcttctc agcggcttta 5760
 tccagcgatt tctattatg tcggcatagt tctcaagatc gacagcctgt caggttaag 5820
 cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880
 caggcagcaa cgtctgtca tngttacaat caacatgcta ccctccgga gatcatcgt 5940
 gtttcaaacc cggcagctta gttgcgctt tccgaatag catcggtaac atgagcaaa 6000
 tctgccgctt tacaacggct ctcccgtga cgcgctccc gactgatggg ctgcctgtat 6060
 cgagtgggta ttttgtgcg agctgcgggt cggggagctg ttggctggct ggtggcagga 6120
 tatattgtgg tgtaacaaa ttgacgtta gacaacttaa taacacattg cggacgtttt 6180
 taatgtactg ggttggtttt ttttttacc agtgagacgg gcaacagctg attgcccttc 6240
 accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgcgc cagcaggcga 6300
 aaatctctgt tgatggtggg toogaaatcg gcaaatccc ttataaatca aaacaatagc 6360
 ccgagatag gttgagtgtt gttccagttt ggaacaagag tcaactatta aagaacgtgg 6420
 actccaacgt caaaggggga aaaaacgtct atcagggcga tggccacta cgtgaacct 6480

cacccaatc aagttttttg gggtcgaggt gccgtaaac actaatcgg aacctaaag 6540

ggagccccc atttagagct tgacggggaa agcggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcggly 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtaacc ggggatcttt cgacactgaa atacgtcgag cctgctcgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gtcatttaag agccagttca tggcggttgg catgatggcc gtcatgcac tgtactcaa 6960

gtacaccaac gctcttttga tccagtcgat catccgctga aggcgccttc gaatctggtt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctc agcgtccctt taaggctgcc 7080

aacagcttcc tcagccagg ccagcccaag accgacaagg cctccctcca gaacccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtgc gaggacggag 7200

cgggtgcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctgacttggg gcattagagg 7320

tcctttggct ttaatatct tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaataca tcgggatgac gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatcgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc agtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

ttcttagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt oggcaaggct acggaagacg gagaagccac cttcagtgga ctgagylacc 7740

atttaattot atttgtgttt gatcgagacc taatacagcc cctacaacga coatcaaagt 7800

ogtagtagta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactatata gaataagata ggtggagagc ttataccgag ctcccaaate 7920

tgteccagatc atggttgacc ggtgcctgga tottccataa gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggg catgacttga gcctaaaate cgcgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaaccg tgactctttc 8100

tggcallycyg agagacggac ggaacgagag agaagggctg agtaataaag caotggccag 8160

acagctcttg cggctctgag gtgcagtgga tgattattaa tccgggaacc gcgcacctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctatt gcctcatcgg 8280

agaatatgga gcttcatoga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcctgccatt aacctaggta cagaagtcca attgcttcg 8400

atctggtaaa agattcacga gatagtaact tctccgaagt aggtagagcg agtacccggc 8460

gcgtaagctc cctaattggc ccataccgga tctgtagggc gtccaaatat cgtgcctctc 8520

ctgcttttgc cgggttatga aaccggaaag gcugctcagg agctggccag cgggcgcgac 8580

cgggaacaca agctggcagt cgaacctacc ggtgctctgc actgacctg ctgaggtccc 8640

tcagtcctgt gtaggcagct ttgcctcgtc tgtccgccg gtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtccea catttgttgc catattttcc tgcctctccc accagctgct 8760

cttttctttt ctctttcttt toccatcttc agtatattca tcttccatc caagaacctt 8820

tatttccct aagtaagtac ttgtctacat ccatactcca tcttcccat cccttattcc 8880

tttgaacctt tcagttcagc ctttccact tcctgcagc ttgactaaca gctacccgc 8940

ttgagcagac atcaccatgc ctgaactcac cgcgagctct gtgcagaagt tctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggogaagaat ctctggcttt 9060
 cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgtaggttt 9120
 ctacaaagat cgttatgttt atcggaactt tgoatcgcc gcgctccga ttccggaagt 9180
 gottgacatt ggggaattca gcgagagcct gacctattgc atctccgcc gtgcacaggg 9240
 tgtcacgttg caagacctgc ctgaaacga actgccgct gttctgcagc cggctcgga 9300
 ggcatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360
 accgcaagga atcggccaat acactacatg gcgtgatttc atatgcgga ttgctgacn 9420
 coatgtgtat cactggcaaa ctgtgatgga cgacacgcgc agtgcgtccg tcgcgcaggc 9480
 tctgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540
 ggatttcggc tccaacaatg toctgacgga caatggccgc ataacagcgg tcattgactg 9600
 gagcaggcgc atgttcgggg attccaata cgaggtgcgc aacatcttct tctgaggcc 9660
 gtggttgct tgtatggagc agcagacgcg ctacttcgag cggaggcacc cggagcttgc 9720
 aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780
 cttggttgac ggcaatttcg atgatgcgc ttgggcgcag ggtcgatgcg acgcaatcgt 9840
 ccgatccgga gccgggactg tggggcgtag acaaatcgcc cgcagaagcg cggccgtctg 9900
 gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcocca gcaactgtcc 9960
 gagggcaaa gaaatagata gatgccgacc gccggatcga tccacttaac gttactgaa 10020
 tcatcaaaaa gcttgacgaa tctggatata agatcgttgg tctgatgto agtcccgag 10080
 ttgagacaaa tgggtttcag gatctcgata agatacgttc atttgtocaa gcagcaaga 10140
 gtgccttcta gtgatttaat agtccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
 cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttoggagac 10320
gagatcaagg agatcaacgg tcgtcaagag acctacgaga ctgaggaato cgctcttgcc 10380
tccacgcgac tatataattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt cgtcaaccag cncctggggt cgcaaagata attgcatgtt 10500
tcttctctga actctcaagg ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctcg 10620
taaccccaa tacgccggcc gaaacttttt tacaactoto ctatgagtcg tttaaccaga 10680
atgcacaggt acactgtgtt agaggtaatc cttctttcta gctagaagtc ctogtgtaact 10740
gtgtaagcgc coactccaca tctccactcg acctgcaggc atgcaagctt gagattaaaa 10800

tagataagga aaagaaagtg aaaagaaatt cgggaagcatg gcacattctt ctttttataa 10860
atacatgctt gacttttttt ttccatcqtat atgatatatg catatgatag atatacaagg 10920
aatctctctc aaggagtttg aaattttgtc ctccaggagc aaaaaaaagt ttttttttat 10980
acatgtttgt acacaagaat agttaaccaat ttgcttttgg cttacgtgct gcaagtttat 11040
atcgttttca atttctttgt ctttacattt tctttgtctt ttatctttcc tcatttagtc 11100
tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcttat tttaactaat 11160
cggcaaacat ccaattttgc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220
ttgattacac totaaatctc gatgaccga caaaaagag cgaacaaga aataatcttg 11280
tgcattcgaa tatgatggaa gattttttcc cctttattct aaatgttgac atagcgtgta 11340
tggtatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt attttatctc 11400
tatgctgctg aagctgcagt caatcagcgt caaggcccg cgcgttgaac tagcccgca 11460

catcacgagg ccacaaagtct gcctgcatgc tcagcggtagc tcgttagttc ggctgcgagt 11520

ggcagcacca cagacagagg aggcgctggg aaccgtgcag gctgccggcg cgggcgatga 11580

gcacagcgcc gatgtagcac tccagcagct tgaccgggct atcgcagagc gtcgtgcccg 11640

gcgcaaacgg gagcagctgt cataccaggc tgcgccatt gcagcatcaa ttggcgtgtc 11700

agccattgcc atcttcgcca cctacotgag atttgccatg cacatgaccg tgggcggcgc 11760

agtgcctagg ggtgaagtgg ctggcactct cctcttggg gttgggtggc cgctcggcat 11820

ggagatgtat gcccgctatg cacacaaagc catctggcat gagtgcctc tgggctggct 11880

gctgcacaag agccaccaca cacctgcac tggaccctt gaagccaacg acttgtttgc 11940

aatcatcaat ggactgcgc ccatgctcct gtgtaccctt ggcttctggc tgcccaacgt 12000

cctggggggc gcctgctttg gagcggggct gggcatcacg ctatacggca tggcatatat 12060

gtttgtacac gatggcctgg tgcacaggcg ctttccacc gggcccatcg ctggcctgcc 12120

ctacatgaag cgcctgacag tggccacca gctacaccac agcggaagt acggtggcgc 12180

gccctggggt atgttcttgg gtccacagga gotgcagcac attccagggt cggcggagga 12240

ggtggagcga ctggctcctg aactggactg gtccaagcgg tagaagcttg agattaaaat 12300

agataaggaa aagaaagtga aaagaaattc ggaagcatgg cacattcttc tttttataaa 12360

tacatgcctg actttctttt tccatcgata tgatatatgc atatgataga tatacaagca 12420

atcttcttca aggagtttga aattttgtcc tccaggagca aaaaaaagt ttttttata 12480

catgtttgta cacaagaata gttaccaatt tgctttggtc ttacgtgctg caagtttata 12540

tcgttttcaa tttctttgtc tttaacattt ctttgtcctt tatctttcct catttagtct 12600

ttgggagaat taggaaaagg gagcggaag gtaagaaatg cttgcgtatt ttactaatc 12660

ggcaaacatc caatttggca aacagcagcc tgtgcaacgc tctcgagatg acagtatctt 12720

tgattacact ctaaattctcg atgaccogac caaaaagagc gaacaaagaa ataattctgt 12780

gcattcgaat atgatgggaag attttttccc cttatttcta aatgttgaca tagcgtgtat 12840

gttatataaa caaaaagaaa ttgtacaaac tttcttttct tctcttttta ttttatctct 12900

atgatccagt tagaacaacc actcagtcac caagcaaaac tgactccagt actgagaagt 12960

aaattctcagt ttaaaggggct tttcattgct attgtcattg ttagcgcatg ggtcattagc 13020

ctgagtttat tactttccct tgacatctca aagctaaaat ttggatggtt attgcctgtt 13080

atactatggc aaacattttt atatacggga ttatttatta catctcatga tgccatgcac 13140

ggcgtagtat ttcccaaaaa caccaagatt aatcatttga ttggaacatt gaccctatcc 13200

ctttatggtc ttttaccata tcaaaaacta ttgaaaaaac attgggttaca ccaccacaat 13260

ccagcaagct caatagaccc ggattttcac aatggtaaac accaaagttt ctttgcttgg 13320

tattttcatt ttatgaaag ttactggagt tgggggcaaa taattgcgtt gactattatt 13380

tataaotttg ctaaatecat actccatctc ccaagtata atctaactta cttttgggtg 13440

ctaccctcgc ttttaagttc attacaatta ttctattttg gtactttttt accccatagt 13500

gaaccaatag ggggttatgt tcagcctcat tgtgccaaa caattagccg tcctatttgg 13560

tggtcattha tcacgtgcta tcattttggc taccacgagg aacatcacga atatctcat 13620

atttcttggg ggcagttacc agaaatttac aaagcaaat agaagcttgg cgtaatcatg 13680

gtcatagctg tttcctgtgt gaaattgta tccgtcaca attccacaca acatacagc 13740

cggagacata aagtgtaaag cctgggggtg ctaatgagtg agctaactca cattaattgc 13800

gltgcgtca ctgcccgtt tocagtcggg aaacctgttg tgccagctgc attaatgaat 13860

cggccaacgc gcggggagag gcggtttgcg tattgggcca aagacaaaag ggcgacattc 13920

aaccgattga gggagggaag gtaaatattg acggaaatta ttcattaaag gtgaattatc 13980

accgtcaccg acttgagcca ttgggaatt agagccagca aaatcaccag tagcaccatt 14040
accattagca aggcgggaaa cgtcaccaat gaaacoatcg atagcagcac cgtaatcagt 14100
agcgaagaa tcaagtttgc ctttagcgtc agactgtagc gcgttttcat cggcatttct 14160
ggctatagcc cctttattag cgtttgcct cttttcataa tcaaaatcac cggaaccaga 14220
gccaccaccg gaacgcctc cctcagagcc gccaccctca gaacgcctac cctcagagcc 14280
accaccctca gagcgcctac cagaaccacc accagagccg ccgcagcagt tgacagagg 14340
ccgatctag taacatagat gacacgcgc gcgataattt atctagttt gcgcgtata 14400
ttttgtttc tatcggtat taaatgtata attgoggac tctaataa aaaaccatc 14460
tcataaataa cgtcatgcat tacatgttaa ttattacatg cttaacgtaa ttcaacagaa 14520
attatatgat aatcatcgca agaccggcaa caggatcca tottaagaaa ctttattgcc 14580
aaatgtttga acgatcggg atcatcggg tctgtggcg gaactccagc aaaatatccg 14640
aacgcagcaa gatatcgcg tgcattctcg tttgctctgg gcagtcgcg ccgagccgt 14700
tgatgtggc gccggggccg atcatattgt cgtcaggat cgtggcgttg tgcctgtcg 14760
ccgttctgt cgtaatgata tggcacctt cgcgcgtc lccgcagag atccgtggg 14820
cgaagaactc cagcatgaga tccccgcgt ggaggatcat ccagccggcg tcccgaaaa 14880
cgattccgaa gcccaacct tcatagaagg cggcggtgga atcgaatct cgtgatggca 14940
ggttggcg cgttggctg gtcatttoga acccagagt ccgcctcaga agaactcgtc 15000
aagaaggcga tagaaggcga tgcgtcoga atcgggagcg gogataccgt aaagcacag 15060
gaagcggtca gccattcgc gcacaagtc ttcagcaata tcaaggtag ccaacgtat 15120
gtctgatag cgttcggca caccagccg gccacgtcg atgaatccag aaaagcggc 15180
atttccacc atgatattcg gcaagcaggc atgcctatg gtcacagca gatcatcgcc 15240

gtcgggcatg cgcgccttga gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc 15300

ttcgtccaga tcactcctgat cgacaagacc ggcttcctac cgagtacgtg ctccgtcctat 15360

gcgatgtttc gcttgggtggc cgaatgggca ggtagccgga tcaagcgtat gcagccgcgc 15420

cattgcatac gccatgatgg atactttctc ggcaggagca aggtgagatg acaggagatc 15480

ctgccccggc acttcgcccc atagcagcca gtcccttccc gcttcagtga caacgtogag 15540

cacagctgcg caaggaacgc ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg 15600

cagttcattc ayygcaucgg acaggtcggc cttgcacaaa agaaccgggc gccctgcgc 15660

tgacagccgg aacacggcgg catoagagca gccgattgtc tgttgtgccc agtcatagcc 15720

gaatagcctc tccaccaag cggccggaga acctcgtgc aatccatctt gttcaatcat 15780

gcgaaacgat ccagatccgg tgcagattat ttggattgag agtgaatatg agactctaata 15840

tggataccga ggggaattta tggaaagtca gtggagcatt tttgacaaga aatatttgc 15900

agctgatagt gaacttaggc gacttttgaa cgcgcaataa tggtttctga cgtatgtgct 15960

tagtcatta aactccagaa acccgcggtc gagtggctcc ttcaacgttg cgtttctgtc 16020

agttccaaac gtaaacggc ttgtcccgcg tcactcggcg ggttcataac gtgactccct 16080

taattctccg ctcatgatca gattgtcgtt tcccgccttc agtttaaac atcagtgttt 16140

gacagatat attgcccggg aaacctaa gaagagagcg tttattagaa taatcggata 16200

tttaaaaggc cgtgaaaagg tttatccgtt cgtccatttg tatgtgcatg ccaaccacag 16260

ggttccccag atctggcgcc ggccagcgag acgagcaaga ttggccgcgc cccgaaacga 16320

tccgacagcg cccccgac aggtgcgcag gcaaattgca ccaacgcata cagccccagc 16380

agaatgccat agtggcggt gacgtcgttc gagtgaacca gatcgcgagc gaggcccgcc 16440

agcaccggca taatcaggcc gatccgaca gcgtcgagcg cgacagtgtc cagaattacg 16500

atcaggggta tgttggtttt cactgtggc ctcggacca gctcgcgtg gtcgattga 16560
acgcgcggat tctttatcac tgataagttg gtggacatat tatgtttatc agtgataaag 16620
tgtcaagcat gacaaagttg cagccgaata cagtgatcgg tgcgcgcctg gacctgttga 16680
acgaggtcgg cgtagacggt ctgacgacac gcaaaactggc ggaacggttg ggggttcagc 16740
agccggcgct ttactggcac ttcaggaaca agcgggcgct gctcgacgca ctggccgaag 16800
ccatgctggc ggaagaatcat acgcattcgg tgcgagagac cgacgacgac tggcgctcat 16860
ttcttgcggc gaattgcgcg agcttcaaggc aggcgctgct cgcctacgcg gatggcgccg 16920
gcatccatgc cggcacgcga ccgggcgcac cgcagatgga aacggccgac gcgcagcttc 16980
gcttcctctg cagggcggtt ttttcggcgg gggacgctt caatgcgctg atgacaatca 17040
gctacttcac tgttggggcc gtgcttgagg agcagggcgg cgacagcgat gccggcgagc 17100
gcggcgccac cgttgaacag gctcgcctct cgcgcctgtt gcgggcgcgg atagacgcct 17160
tcgacgaagc cggtcgggac gcagcgttcg agcagggact cgcggtgatt gtcgatggat 17220
tggcgaaaag gaggtctggt gtcaggaaag ttgaaggacc gagaaaggtg gacgattgat 17280
caggacgcgt gccggagcgc aaacccactca ctacgcgcga ggcattgtaga caacatcccc 17340
tcccccttcc caccgcgtca gacgccgta gcagccgcct acgggctttt tcattgccctg 17400
ccctagcgtc caagcctcac ggcgcgcgtc ggcctctctg gcggccttct ggcgctcttc 17460
cgcttctctg ctcactgact cgtgcgcgtc ggtcgttcgg ctgcggcgag cggtatcagc 17520
tcactcaaa ggcgtaatac ggttatccac agaatacggg gataacgcag gaaagaacat 17580
gtgagcaaaa ggccagcaaa aggcacggaa ccgtaaaaag gccgcgttgc tggcggtttt 17640
ccataggctc cgcgccctcg acgagcatca caaaaatcga cgtcaagtc agaggtggcg 17700
aaacccgaca ggactataaa gataccaggc gtttccctt gyaagctccc tcgtgcgctc 17760

tcctgttccg accctgcgcg ttaccggata cctgtccgc tttctccctt cggsaagcgt 17820
ggcgcttttc cgtgcataa cctgcttcg gggtcattat agcgattttt tcggatatatc 17880
catccttttt cgcacgatat acaggatttt gccaaagggg tcgtgtagac ttctcttggt 17940
gtatccaacg cgcgcagcgc ggcaggatag gtgaagtagg cccacccgcg agcgggtggt 18000
ccttcttcac tgtcccttat tcgcacctgg cgggtgcctaa cgggaatcct gctctgcgag 18060
gctgscggcg taccgcggcg gtaacagatg agggcaagcg gatggctgat gaaaccaagc 18120
caaccaggaa gggcagccca cctatcaagg tgtactgctt tccagaogaa cgaagagcga 18180
ttgagggaaa ggcggcgggc gccggcatga gcctgtcggc ctacctgctg gcgctgggcc 18240
agggctacaa aatcacggcg gtcgtggact atgagcacgt ccgcgagctg gcccgcatca 18300
atggcgacct gggcgcgctg ggcggcctgc tgaaactctg gctcaccgac gaccgcgca 18360
cggcgcggtt cgggtgatgc acgatcctcg cctgtcggc gaagatcgaa gagaagcagg 18420
acgagcttgg caaggtcatg atgggcgtgg tccgcccag ggcagagcca tgactttttt 18480

agccgctaaa acggccgggg ggtgcgcgtg attgccagc acgtcccat gcgctccatc 18540
aagaagcgcg acttcgcgga gctgggtgaag tacatcacgc acgagcaagg caagaccgag 18600
cgcttttcg acgtca 18617

<210> 51

<211> 18333

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>
 <221> misc_feature
 <222> (10264)..(10264)
 <223> n i s a, c, g, o r t

<220>
 <221> misc_feature
 <222> (10472)..(10472)
 <223> n i s a, c, g, o r t

<220>
 <221> misc_feature
 <222> (10563)..(10563)
 <223> n i s a, c, g, o r t

<400> 51
 cggggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgcag 60
 aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctgcggg 120
 aaaacttggc cctcactgac agatgagggg cggacgttga caattgaggg gccgactcac 180
 cccgcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240
 cagcctcgca aatcggcgaa aacgcctgat ttacgcgag ttcccacag atgatgtgga 300
 caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360
 gaggggcgcg atccttgaca cttgaggggc agagtgtctga cagatgaggg gcgcacctat 420
 tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgccggt 480
 ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg 540
 tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600
 cttctogaac cctccgggac cgttaacggg ggcctcccat ccccccaggg gctgcgcccc 660
 tcggccgcga acggcctcac ccaaaaatg gcagcgctgg cagtccttgc cattgccggg 720
 atcggggcag taacgggatg ggcgatcagc ccgagcgga cggccggaag cattgacgtg 780

ccgcagggtgc tggcatcgac attcagcgac cagggtgccg gcagtgaggg cggcggcctg 840

gggtggcgcc tgccttcac ttggccgctc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgccg gtccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatatgac aatactgata 1140

agataatata tcttttatat agaagatctc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgcttg aaaccaggga caataacott atagcttgta aattctatca taattgggta 1320

atgactccaa ctatttgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgcgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcgccag ccatcgtca tccatctac cactcaaaag 1560

ggtgacagca ggtcctaag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaacgg tcttcggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttctgtccatt tccgcgcaga cgatgacgtc actgcccgc 1740

tgatgcgcg aggttaacga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatcgcg gctgttgccc ggcattcaac gccattcatg gccatatcaa 1860

tgatttttctg gtgngtacgg ggttgagaag cgggtgaagt gaactgcagt tcccatgttt 1920

tacggcagtg agagcagaga tagcgtgat gtccggcggt gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100

aaatcggtc cgtcgatact atgttatcgc ccaactttga aaacaacttt gaaaagctg 2160

ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggaagtc gtcttgttat 2220

aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatacccg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaaggatata taagctgggt ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaggga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtctctgact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gggagtgca tcaggtcttt 2640

tactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaga 2820

ggaacttgtc ttttccacg gcyacutggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgac ttgggagaag cggcagggcg gacaagtggc atgacattgc 2940

cttctcgctc cgtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatatct tactggatga 3060

attgttttag tacctagatg tggcgcaacg atccggcgca caagcaggag cgcacogact 3120

tcctccgcat caagtgtttt ggctctcagg ccgagggcca cggcaagtat ttgggcaagg 3180

ggctcgctgg attcgtgcag ggcaagatc ggaataccaa gtacgagaag gacggccaga 3240

cggctacgg gacogacttc attgcgata aggtggatta tctggacacc aaggaaccag 3300

gcgggtcaaa tcaggataaa gggcacattg ccccgcggtg agtcggggca atccccgaag 3360

gagggtgaat gaatcggacg tttagccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttcgcg cgaggatgcc gaaaccatcg caagccgcac cgtcatgctg gcgccccgcg 3480

aaacgttcca gtccgtcggc tcgatgggtc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cggcatcggc cggcgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaaacta 3660

tgcagacuaa yaaycyaaaa accgcggcgg aggaacctgg aaaacaggto agcgaggcca 3720

agcaggccgc gttgtctgaa cacacgaagc agcagatcaa ggaatgcag ctttccctgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cgcgcgcaac aagaaaatcc cgcgcgaggg gctgcaaaac aaggtcattt 3900

tcacgttcaa caaggacgtg aagatcaact acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagcoga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat gcccggtatt 4080

aacgaaggc cgaggaaatc ctgtcgcgcc taacggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcgggtgcgc tgetgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggctctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaattc atatgggaga agtaccccaa gctgtcgccg acggcccgac 4320

ggatgttcga ctatttcagc tcgcacgggg agccgtaccc gctcaagctg gaaaccttc 4380

gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtccggogaag 4440

cctgcgaaga gttcgcaggc agcggcctgg tggaacacgc ctgggtcaat gatgaacctg 4500

tgcattgcga acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtc 4620
 gctcgggacg caccggcgcg tctacgaact gccgataaac agaggattaa aattgacaat 4680
 tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740
 ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
 gaaaagccc atggaggcgt tcgtgaacg gttgcgagat gccgtggcat tcggcgccca 4860
 catcgacggc gagatcattg ggtgttcggt ctccaacag gaggaaggcc ccaaggacgc 4920
 tcacaaggcg catctgltuy gcgttttcgt ggggcgcgaa cagcgaggcc gagggttcgc 4980
 cggtatgtcg ctgcgggcgt tgcggcggg tttattgctc gtgatgatcg tccgacagat 5040
 tccaacggga atctggttga tgcgcattct catctcggc gcaactaata ttctgcattt 5100
 ctggagcttg ttgtttattt cggctctacc cctgccgggc ggggttcggc gcacggtagg 5160
 cgtgtgcag ccgctgatgg tcgtgttcct ctctgcgcct ctgctaggta gcccgatacg 5220
 attgatggcg gtctcggggc ctatttcggc aactgcgggc gtggcgtgtg tgggtttgac 5280
 accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
 ggcgttcgga acogtgctga cccgcaagtg guaaculccc gtgcctctgc tcacctttac 5400
 cgctcggcaa ctgcgggcgc gaggacttct gctcgttcca gtacgtttag tgtttgatcc 5460
 gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gctgatcgg 5520
 agcgggttta acctactcc tttggttcgc ggggatctcg cgaactcgaac ctacgttgt 5580
 ttcttactg ggttttctca gccccagatc tggggtcgat cagccgggga tgcacaggc 5640
 cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgcgcaatgg ataggggagt 5700
 tgatatcgto aacgttcaact tctaaagaaa tagcgccact cagcttctc agcgggttta 5760
 tccagcgatt tctattatg tcggcatagt tctcaagatc gacagcctgt cagggttacy 5820

cgagaaatga ataagaaggc tgataattcg gatctctcg aggagatga tatttgatca 5880
 caggcagcaa cgetctgtca tcgttacaat caacatgcta cctccgcga gatcatccgt 5940
 gtttcaaaacc cggcagctta gttgcogtto ttccgaatag catcggtaac atgagcaaag 6000
 tctgccgcct tacaacggct ctcccgctga cgcogtcccg gactgatggg ctgcctgtat 6060
 cgagtgggtga ttttgtgccg agctgccggg cggggagctg ttggctggct ggtggcagga 6120
 tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
 taatgtactg ggggtggttt lullllcacc agtgagacgg gcaacagctg attgccttc 6240
 accgcctggc cctgagagag ttgcagcaag cggtccacgc tggtttgccc cagcaggcga 6300
 aaatcctgtt tgatgggtgt tccgaaatcg gcaaaatccc ttataaatca aagaatagc 6360
 ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaactgtg 6420
 actccaaagt caaagggcga aaaacogtct atcagggcga tggccacta cgtgaacct 6480
 cacccaaac aagttttttt gggtcgaggt gccgtaaag actaaatcgg aaccctaaag 6540
 ggagccccc atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaagga 6600
 agaaagcgaa aggagcggg gccattcagg ctgcgcaact gttgggaagg gcgatcggg 6660
 cgggcctctt cgtattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
 tgggtaacgc cagggttttt ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780
 ctccgtatcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840
 cgaggagcct cgtcctgtca caactacaa catggagtac gataagggcc agttccacca 6900
 gctcattaag agccagttca tgggcgttgg catgatggc gtcatgcac tgtaactcaa 6960
 gtacaccaac gctcttctga tccagtcgat catccgctga aggcgcttcc gaatctggtt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080
aacagctttc tcagocaggg ccagcccaag accgacaagg cctccctoca gaacgcgag 7140
aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggay 7200
cgggtgcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagratagct tcatttcgat ttgctttcca ggctgagact ctgacttga gcatagaggg 7320
tcctttggct ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt 7440
tcgaaataca tccggatgtc gaaggcttgg ggcacctcgc ttggttgaat ttagaacgtg 7500
gcactattga tcattccgata gctctgcaaa gggcggttga caatgcaagt caaacgttgc 7560
tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620
tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680
gtaaggattt cggcacgggt acggaagacg qagaagccac cttoagtga ctgcagtacc 7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga coatcaaatg 7800
cgtagtagta ccagtggagga agtggaacta aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aaactatata gaataagata ggtggagagc ttataccgag ctoccaaact 7920
tgtccagatc atggttgacc ggtgcctgga tcttctata gaatcatcct tattcgttga 7980
cctagctgat tctggagtga ccagaggggt catgacttga gctaaaact cgcgcctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaacgg tgactctttc 8100
tggcatgcgg agagacggac ggaagcagag agaagggctg agtaataac cactggccag 8160
acagctctgg cggctotgag gtgcagtgga tgattattaa tccgggacgg gccgccctc 8220
cgccccgaag tggaaaggct ggtgtgcccc togttgacca agaattctatt gcacatcgg 8280

agaatatgga gcttcatoga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
 gtatagcogt oggogaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400
 atctggtaaa agattcacga gatagtacct tctcogaagt aggtagagcg agtaccggc 8460
 gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
 ctgctttgoc oggtgtatga aacoggaaag gcogetcagg agctggccag oggogcagac 8580
 cgggaacaca agctggcagt cgaacctoc ggtgctctgc actcgacctg ctgaggctcc 8640
 tcagtcctg gtaggcagct ttgcccgtc tgtccggccg gtgtgtcggc ggggttgaca 8700
 aggtcgttgc gtcagtcocaa catttgttgc catattttcc tgcctctccc accagctgct 8760
 cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
 tatttcctt aagtaagtac ttgtctacat ccatactcca tcttcccat ccttatttcc 8880
 tttagaacctt tcagttcgag ctttccact tcacgcagc ttgactaaca gctacccgc 8940
 ttgagcagac atcaccatgc ctgaactcac cgcgactct gtcgagaagt ttctgatcga 9000
 aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctgctgcttt 9060
 cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgagtgttt 9120
 ctacaaagat cgttatgttt atcggcactt tgcacggcc gogctccga ttccggaagt 9180
 gcttgacatt ggggaattca gcgagagcct gacctattgc atctccgcc gtgcacaggg 9240
 tgtcacgttg caagacctgc ctgaaaccga actgcccgt gttctgcage cggctcgga 9300
 ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360
 acggcaagga atcggtcatt acactacatg gcgtgatctt atatgngcga ttgctgaccc 9420
 ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcttccg tcgcgcaggc 9480
 tctgatgag ctgatgcttt gggccgagga ctgcccga gtcgggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tctgaacga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttgg atgatgcagc ttggggcgag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggcgcgtcg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccc aactcgttc 9960

gagggcacaag gaatagagta gatgcgcacc gggggatcga tccacttaac gttactgaaa 10020

tcaccaaaca gcttgacgaa tctggatata agatcgttgg tgtcgatgto agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaaga 10140

gtgccttota gtgatttaat agctccatgt caacaagaat aaacgcggtt ttggggttta 10200

cctcttccag atacagctca tctccaatgc attaatgcat tgactgcac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tttttcatt ttccgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag aactacgaga ctgaggaatc cgtctcttgc 10380

tcacgcgcac tatatatttg tctctaattg tactttgaca tgcctctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctggggt gcgaaagata attgcattgt 10500

ttctccttga actctcaagc ctacaggaca cacattcatc gtagggtataa acctcgaaat 10560

canttctac taagatggta tacaatagta accatgcatt gttgcotagt gaatgctccg 10620

taacacccaa tacgcgggcc gaaacttttt tacaactctc ctatgagtgc tttaaccaga 10680

atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctctgttact 10740

gtgtaagcgc ccactccaca tctccactgc acctgcaggc atgcaagctt gagattaaaa 10800

tagataagga aaagaagtg aaaagaaatt cggaagcatg gcacattctt cttttataa 10860

atacatgcct gacttttctt ttccatcgat atgatatatg catatgatag atatacaagc 10920

aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaagt tttttttat 10980

acatgtttgt acacaagaat agttaccaat ttgcttttgt cttacgtgct gcaagtttat 11040

atogttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100

tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgtat tttactaatt 11160

cggcaaacat ccaatlltgc aaacagcagc ctgtgcaacg ctctcgagat gacegtatct 11220

ttgattacac tctaaatctc gatgaoccca ccaaaagag cgaacaaaga aataatcttg 11280

tgcattcgaa tatgatggaa gattttttcc cccttattct aaatgttgac atagcgtgta 11340

tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt attttatctc 11400

tatgtttgtg atttggaaat ccctgatcgt ttctgttacc gtgattggca tgggaagtgt 11460

tgctgcactg gcacacaaat acatcatgca cggctggggg tggggatggc atctttcaca 11520

tcatgaacgg cgtaaagggt cgtttgaagt taacgatctt tatgccgtgg tttttgctgc 11580

attatogac ctgctgattt atctgggcag tacaggaaat tggccgctcc agtggattgy 11640

cgaggtatg acggcgtatg gattactcta ttttatgggt caccgacggg tgggtgatca 11700

acgttggcca ttccgtata ttccacgcaa gggctacctc aaacggttgt atatggcgca 11760

ccgatgcat caccgcgtca ggggcaaaga aggttggtt tcttttggct tctctatgc 11820

gcgcgccctg tcaaaacttc aggcagcgt cggggaaga catggcgcta gagggggcgc 11880

tgccagagat gcgcaggggc gggaggatga gccgcgcatc gggaagttag ggccgtacca 11940

gaggcggcca gcagcagcgt taatttttgc ggcgtggtcg ttgactgcgc ctgatcccaa 12000

agcttgagat taaaatagat aaggaaaaga aagtgaaaag aaattcggaa gcattggcaca 12060

tttttttttt tataaataca tgcctgactt tttttttcca tcgatgatgat atatgcatat 12120

gatagatata caagcaatct tcttcaagga gtttgaaatt ttgtcctcca ggagcaaaaa 12180

aaagtttttt tttatacatg tttgtacaca agaatagtta ccaatttgct ttggtcttac 12240

gtgctgcaag tttatatcgt tttcaatttc ttgtcttta ctttttcttt gtccctttatc 12300

tttctcatt tagtctttgg gagaattagg aaaagggagc ggaaaggtaa gaaatgcttg 12360

cgtattttac taattcggca aacatccaat ttggcaaca gcagcctgtg caacgccttc 12420

gagagacag tatctttgat tacaacttaa atctogatga ccgaacaaa aagagcgaa 12480

aaagaaataa tcttgtgcatt tcgaatatga tggaagattt tttccctctt attctaaatg 12540

ttgacatagc gtgtatgtta tataaaca aaagaattgt acaaaccttc ttttcttctc 12600

tttttatttt atctctatga tccagttaga acaaccactc agtcatcaag caaaactgac 12660

tcagactagc agaagtaaat ctcaagttta ggggcttttc attgctattg tcattgttag 12720

cgcattgggc attagcctga gtttattact ttcccttgac atctcaaagc taaaattttg 12780

gatgttattg cctgttatac tatggcaaac atttttatat acgggattat ttattacatc 12840

tcattgatgc atgcattggc tagtatttcc ccaaacac aaatttattc atttgatttg 12900

aacattgacc ctatcccttt atggtctttt accatatcaa aaactattga aaaacattg 12960

gttacaccac cacaatccag caagctcaat agaccggat ttccacaatg gtaaacacca 13020

aagtctcttt gcttggtatt ttcatcttat gaaaggttac tggagttggg ggcaataaat 13080

tgcgttgact attatttata actttgctaa atacatactc catatcccaa gtgataatct 13140

aacttacttt tgggtgctac cctcgtcttt aagttcatla caattattct attttggatc 13200

ttttttacc catagtgaac caataggggg ttatgttcag cctcattgtg cccaacaat 13260

tagccgtcct atttggtggt catttatcac gtgctatcat ttggctacc acgaggaaca 13320

tcacgaatat cctcatattt cttggtggca gttaccagaa atttacaag caaaatagaa 13380

gcttggcgta atcatggta tagctgttcc ctgtgtgaaa ttgttatccg ctcaaatcc 13440

cacacaacat acgagccgga agcataaagt gtaaacctg gggatgctaa tgagtgaagct 13500

aactcacatt aattgcgttg cgtcactgc cgcctttcca gtcgggaac ctgctgtcc 13560

agctgcatta atgaatcggc caacgcgcgg gagagggcg tttgcgtatt gggcaaga 13620

caaaaggcg acattcaacc gattgaggga gggaaggtaa atattgacgg aaattattca 13680

ttaaaggta attatccgc tcaccgctt gagcatttg ggaattagag ccagcaaat 13740

caccagtagc accattacca ttgcaaggc cggaaacgtc accaatgaaa ccatgatag 13800

cagcaccgta atcagtagcg acagaatcaa gtttgcttt agcgtcagac tgtagcgt 13860

tttcatcggc attttcggto atagccctt tattagcgtt tgcatcttt tcataatcaa 13920

aatcaccgga accagagcca ccaccggaac cgcctcctc agagccgcca cctcagaac 13980

cgccaccctc agagccacca cctcagagc cgccaccaga accaccacca gagccgcgc 14040

cagcattgac aggaggccgc atctagtaac atagatgaca ccgcgcgcga taatttatcc 14100

tagtttgccg gctatatttt gttttctatc gcgtattaaa tgtataattg cgggactcta 14160

atcataaaaa cctatctcat aaataacgtc atgcattaca tgtaattat tacatgctta 14220

acgtaattca acagaaatta tatgataac atcgcaagac cggcaacagg attcaatctt 14280

aagaactttt attgccaat gtttgaacga tcggggatca tccgggtctg tggggggaac 14340

tcaccgaaaa tatoogaaag cagcaagata tcggggtgca tctcgtctt gcttgggag 14400

tcgcccga cgcctgtgat gtggacgcgg ggcccgatca tattgtcgtc caggatcgtg 14460

gcgtgtgtct tgcgcgcgt tgctgtcgta atgatatcgg cacttcgac cgcctgttcc 14520

gcagagatcc cgtggcgcaa gaactccagc atgagatccc cgcgctggag gatcatccag 14580
ccggcgctccc ggaaaacgat tcogaagccc aacctttcat agaaggcggc ggtggaatcg 14640
aaatctcgtg atggcagggtt gggcgctcgt tggtcggtca ttctgaaccc cagagtcgccg 14700
ctcagaagaa ctctcaaga aggcgataga aggcgatcgc ctgcgaatcg ggagcggcga 14760
taccgtaaag cacgaggaag cggtcagccc attcgccgcc aagctcttca gcaatatcac 14820
gggtagccaa cgtatgtcc tgatagcggc ccgccacacc cagccggcca cagtcgatga 14880
atccagaaaa gcggccattt tccaccatga tattoggcaa gcaggcatcg ccatgggtca 14940
cgacgagatc atcgccgtcg ggcgatcgcg ccttgagcct ggogaacagt tcggctggcg 15000
cgagcccctg atgctcttcg tccagatcat cctgatcgac aagaccggct tccatccgag 15060
tacgtgctcg ctgatcgca tgtttcgctt ggtggtcgaa tgggcaggta gccggatcaa 15120
gcgatgcag ccgcgcgatt gcacacagca tgatggatac ttctctggca ggagcaaggt 15180
gagatgacag gagatcctcc cccgcactt cgcacaatag cagccagtcc ctctccgctt 15240
cagtgacaac gtgcagcaca gctgcgcaag gaacgcccg cgtggccagc cacgatagcc 15300
gcgctgcctc gtctcgcagt tcattcaggg caccggacag gtcggtcttg acaaaaagaa 15360
ccggcgcccc ctgcgctgac agccggaaca cggcgccatc agagcagcgc attgtctgtt 15420
gtgccagtc atagocgaat agcctctcca cccaagcggc cggagaacct gcgtgcaatc 15480
catcttgttc aatcatcgca aacgatccag atccggtgca gattabttgg attgagagtg 15540
aatatgagac tctaattgga taccgagggg aatttatgga acgtcagtgg agcatttttg 15600
acaagaaata tttctagct gtagtgacc ttaggcgact ttggaacgc caataatggt 15660
ttctgacgta tgtgcttagc tcattaaact ccagaaaccc cgggctgagt ggctccttca 15720
acgttcggtt tctgtcagtt ccaaacgtaa aacggcttgt cccgcgtcat cggcgggggt 15780

cataacgtga ctcccttaatt tctccgctca tgatcagatt gtcgtttccc gccttcagtt 15840
taaactatca gtgttttgaca ggatatattg gcgggttaac ctaagagaaa agagcgttta 15900
ttagaataat cggatatttta aaagggcgtg aaaagggtta tccgttcgtc catttgtatg 15960
tgcatgccaa ccacagggtt cccagatctt ggcgcggcc agcgagacga gcaagattgg 16020
ccccccccc aaacgatccg acagcgcgc cagcacaggc gcgcaggcaa attgcacca 16080
cgcatacagc gccacgagaa tgccatagtg ggcggtgacg tcgttcgagt gaaccagatc 16140
gcgcaggagg cccggcagca ccggcataat caggccgatg ccgacagcgt cgaagcgcac 16200
agtgtctaga attacgatca ggggtatgtt ggggttcacg tctggcctcc ggaccagcct 16260
ccgtgggtcc gattgaacgc gcggattctt tatcactgat aagtgggtgg acatattatg 16320
tttatcagtg ataaagtgtc aagcatgaca aagttgcagc cgaatacagt gatccgtgcc 16380
gcccctggacc tgttgaacga ggtcggcgtg gacggtctga cgacacgcaa actggcggaa 16440
cggttggggg ttacgcagcc gccgcctttac tggcacttca qqaacaagcc ggcgctgctc 16500
gacgcactgg ccgaagccat gctggcggag aatcatacgc attcgggtgcc gagagccgac 16560
gacgactggc gctcatttct gatcgggaat gccgcgagct tcaggcaggc gctgctcgcc 16620
taccgcgatg gcgcgcgcac ccacgcgcgc acgcgacccg gcgcacgcga gatggaaacg 16680
gcgcacgcgc agcttcgctt cctctcgagc gcgggttttt cggccgggga cgcctcaat 16740
gcgctgatga caatcagcta cttcactgtt ggggccgtgc ttgaggagca ggcggcgac 16800
agcgatgcgc gcgagcgcgc gcgcacggtt gaacaggctc cgctctcgcc gctgttcgcy 16860
gcgcgatag acgccttoga cgaagccggt ccggacgcag cgttcgagca gggactcgcg 16920
gtgattgtcg atggattggc gaaaaggagg ctcgttgtca ggaacgttga aggaccgaga 16980
aagggtgacg attgatcagg accgctgcgc gagcgcaacc cactcactac agcagagcca 17040

tgtagacaac atccctccccc cctttccacc gcgtcagacg ccgctagcag cccgctacgg 17100
gctttttcat gccctgccct agcgtccaaag cctcaaggcc gcgctcgcc tctctggcgg 17160
ccttctggcg ctcttcogct tectogctca ctgactcgct gcgctcggtc gttcggtctc 17220
ggcgagcggt atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata 17280
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaacct aaaaaggccg 17340
cgttgctggc gtttttccat aggtctccgc cccctgacga gcatacaaaa aatcgacgct 17400
caagtacagag gtggcgaaac ccgacaggac tataaagata ccaggcggtt cccctggaa 17460
gtccctcgct gcgctctctt gttccgacct tgccgcttac cggataacct tccgccttcc 17520
tccttcggg aagcgtggcg cttttccgct gcataacct gcttcggggg cattatagcg 17580
atthtttcgg tatatccatc ctttttcgca cgatatacag gattttgcca aagggttcgt 17640
gtagactttc cttggtgtat ccaacggcgt cagccgggca ggataggtga agtagggcca 17700
cccgccacgc qqtgttctt cttcactgct ccttattcgc acctggcggt gctcaacggg 17760
aatcctgctc tgcgaggctg gccggtacc gccggcgtaa cagatgaggg caagcggatg 17820
gctgatgaaa ccaagccaac cagggaaggc agccaccta tcaaggtgta ctgccttcca 17880
gacgaacgaa gagcgattga ggaaaaggcg gcggcgcccg gcgatgacct gtcggcctac 17940
ctgctggcgg tcggccaggg ctacaaaatc acgggcgtcg tggactatga gcacgtccgc 18000
gagctggccc gcataaatg cgacctgggc cgcctgggcs gcctgctgaa actctggctc 18060
accgacgacc cgcgcacggc gccgttcggt gatgccacga tectcgccct gctggcgaa 18120
atcgaagaga agcaggacga gcttgcaag gtoatgatgg gcgtggtccg cccgaggcca 18180
gagccatgac ttttttagcc gctaaaacgg cgggggggtg cgcgtgattg ccaagcagt 18240
cccatcgcc tccatcaaga agagcgactt cgcggagctg gtgaagtaca tcaccgacga 18300

gcaaggcaag accgagcgcc ttgcgacgc tca

18333

<210> 52

<211> 17

<212> DNA

<213> Artificial

<220>

<223> Primer

<220>

<221> misc_feature

<222> (3)..(3)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (9)..(9)

<223> n is a, c, g, or t

<400> 52

gcngarggna thtggtta

17

<210> 53

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Primer

<220>

<221> misc_feature

<222> (3)..(3)

<223> n is a, c, g, or t

<220>

<221> misc_feature
<222> (6)..(6)
<223> n is a, c, g, or t

<400> 53
tongcnagra adatrtrtg

20

<210> 54
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 54
aagtgacacc ggttacacgc ttgtctt

27

<210> 55
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 55
gcttatcacc atctgttaac tocttgc

27

<210> 56
<211> 32
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 56
agagagggat ccttaaatgc gaatatogtt gc

32

<210> 57
<211> 32
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 57
agagaggggat ccatgtctga tcaaaagaag ca

32

<210> 58
<211> 37
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 58
actttattgg atccttaaat gogaatatcg ttgctgc

37

<210> 59
<211> 38
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 59
gttccaattg gccacatgaa gagtaagaca ggaacag

38

<210> 60
<211> 38
<212> DNA
<213> Artificial

<220>

<223> Primer

<400> 60

cctgtctttac tcttcatgtg gccaatggg accaacaac

38

<210> 61

<211> 38

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 61

ctatttttaac catatgtctg atcaaaagaa gcatattg

38

<210> 62

<211> 16103

<212> DNA

<213> Artificial

<220>

<223> Primer

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 62

gatcttttoga cactgaata cgtcagacct gtcgcttg gaagcggcga ggagcctcgt	60
cctgtcaca ctaccaacat ggagtacgat aagggccagt tocgccagct cattaagagc	120
cagttcatgg cgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct	180
cttctgatcc agtcgatcat ccgctgaagg cgttttcgaa tctggttaag atccacgtct	240
tccgggaagcc agcgactggt gacctccagc gtccctttaa ggctgccaac agctttctca	300
gccagggccca gcccaagacc gacaaggccl cccctccagaa cgcgcgagag aactggaggg	360
gtgggtgtcaa ggaggagtaa gtcctctatt gaagtcggag gacggagcgg tgtcaagagg	420
atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca	480
tttgatttg ctttcaggc tgagactcta gcttggagca tagagggtcc tttggctttc	540
aatattctca agtatctoga gtttgaactt attcctgtg aaccttttat tcaccaatga	600
gcattggaat gaacatgaat ctgaggactg caatgcgcat gagggttttcg aaatacatcc	660
ggatgtcgaa ggcttggggc acctgcgtg gttgaattta gaacgtggca ctattgatca	720
tccgatagct ctgcaaaggc cgttgacaaa tgcaagtc aaagtgtcag cagttccagg	780
tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct	840
caccacaaaa gtcagacggc gtaacaaaa gtcacacaa acaaagctga aggatttcgg	900
caaggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt	960
tgtgtttgat cgagacctaa tacagccctt acaacgacca tcaaagtcgt atagctacca	1020
gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa	1080
ctatacagaa taagataggt ggagagctta taccgagctc ccaaattcgt ccagatcatg	1140
gttgaccggt gcctggatct tcctatagaa tcctccttat togttgacct agctgattct	1200

ggagtgaccc agaggggtcat gacttgagcc taaaatccgc cgcctccacc atttgtagaa 1260

aaatgtgacg aactcgtgag ctctgtacag tgaccygtga ctctttctgg catgoggaga 1320

gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380

ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440

aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcacgggaga atatggagct 1500

tcacgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagcgcctcg 1560

cgaatatgca tgcattaac ctaggctacg aagtccaatt gcttccgac tggtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680

aattggccca tcggcatct gtggggctc caaatatcgt gcctctcctg ctttgcccg 1740

tgatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcact cgacctctg aggtccctca gtcctcggt 1860

ggcagcttg cccgctctgt ccccccgtg tgcggcggg gttgacaagg tcgttcgctc 1920

agtcacaat ttgttgcat attttctgc tctcccacc agctgctctt tctttttctc 1980

tttttttcc catcttcagt atattcatct tcccatcaa gaacctttat ttccttaag 2040

taagtacttt gctacatcca tactcatcc ttccatccc ttattcttt gaacctttca 2100

gttcgagctt tccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160

accatgcctg aactaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220

gtctccgacc tgatgcagct ctggagggc gaagaatctc gtgctttcag cttcgatgta 2280

ggggggcgtg gatattgct cgggtaaat agctgcgccg atggtttcta caaagatcgt 2340

tatgtttatc ggcactttgc atcggccgcg ctcccgttc cggaagtgtc tgacattggg 2400

gaattcagcg agagcctgac ctattgcatc tcccgcctg cacagggtgt caggttgcaa 2460

gacctgcttg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgog 2520
 atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaagggaatc 2580
 ggatcaataga ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640
 tggcaaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctctc cgatgagctg 2700
 atgcttttggg ccgaggactg cccgaagtc cggcacctcg' tgcacggga tttcggtctc 2760
 aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820
 ttccggggatt cccaatacga ggtccgcaac atcttcttct ggaggccgtg gttggcttgt 2880
 atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgacgg atcgcccgcg 2940
 ctccggggcgt atatgctccg cattgggtctt gaccaactct atcagagctt ggttgacggc 3000
 aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgctcg atccggagcc 3060
 gggactgtcg ggcgtacaca aatgcgccg agaagcgcg ccgtctggac cgatggctgt 3120
 gtagaagtac tcgcgatag tggaaaccga cgcgccagca ctgctccgag ggcaaggaa 3180
 tagagtagat gccgaccgcg ggcgcgatcc acttaacgtt actgaaatca tcaaacagct 3240
 tgacgaatct gcatataaga tcgttggtgt cgatgtcagc tcgggagttg agacaaatgy 3300
 tgttcaggat ctgcataaga tacgttcatt tgtccaagca gcaaagagtg ccttctagt 3360
 atttaatatg tcatgtcaa caagaataaa acgcgttttc ggggttacct ctccagata 3420
 cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgctt ncaggctccg 3480
 gcgaagagaa gaatagetta gcagagctat ttctatttcc gggagacgag atcaagcaga 3540
 tcaacggctg tcaagagacc tacgagactg aggaatccgc tcttggtctc acgcgactat 3600
 atatttgtct ctaattgtac ttgacatgc tctcttctt tactctgata gcttgactat 3660
 gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcattgttct tcttgaact 3720

ctcaagccta caggacacac attcatogta ggtataaacc tcgaaatcan ttccactaa 3780
gatggtatac aatagtaacc atgoatgggt goctagttaa tgctccgtaa caccocaatc 3840
gcggcgcgaa acttttttac aactctccta tgagtogttt aocagaaatg cacaggtaca 3900
cttgtttaga ggtaatcctt ctttctagct agaagtcctt gtgtactgtg taagcgccca 3960
ctccacatct ccaactcgacc tgcaggcatg caagcttgag tctatcgctt ccaaaaagta 4020
cgggtgctgaa ttcagatata aatcgctgtg tgctaaaatt aacactgtcg ataaagacaa 4080
gcgtgtaacc ggtgtcactt tggaaagcgg agaaagcatt gaagccgatg cagtcgtatg 4140
taatcgcatc cttgtttatg cttatcacca tctgttacct ctttgcaatt ggacaagaa 4200
gacattagcc tcaagaanaa tcaacttcac atctatttcg ttttattggt ccatgtcaac 4260
aaaggtgctt caattagacg tacacaatat cttcttggtt gaagcctaca aggaagttt 4320
tgatgagatt ttcaacgact tgggtttgoc ctctgaagct tggcgtaac atggtcatag 4380
ctgtttcctg tgtgaaattg ttatcgcctc acaattccac acaacatag agccggaagc 4440
ataaagtgta aagcctgggg tgctaatga gtgagctaac tcacattaat tgcgttgcc 4500
tcaactgccc ctttccagtc gggaaacctg tctgtccagc tgcattaatg aatcggccaa 4560
cgcgcgggga gaggcgtttt gctattggg ccaagacaa aaggcgacaa ttcaaccgat 4620
tgaggggagg aaggtaataa ttgacggaaa ttattcatta aaggtgaatt atcaccgtca 4680
ccgacttgag ccatttggga attagagcaa gcaaatcac cagtagcaac attaccatta 4740
gcaaggccgg aaagctcacc aatgaaacca tggatagcag caccgtaac agtagcgaca 4800
gaatcaagtt tgcccttagc gtcagactgt agcgggtttt catcgccatt ttccgtcata 4860
gccccttat tagcgtttgc catcttttca taatcaaat caccggaacc agagccacca 4920
ocggaacgc ctcctcaga gcggcaccc tcagaaccgc caccctcaga gccaccacc 4980

tcagagccgc caccagaacc accaccagag ccgccgccag cattgacagg aggcccgatc 5040

tagtaacata gatgacacgg ccgccgataa ttatcctag ttgcgcgct atattttgtt 5100

ttctatcgcg tattaatgt ataattgcgg gactctaabc ataaaaacc atetcataaa 5160

taacgtcatg cattacatgt taattattac atgcttaacg taattcaaca gaaattatat 5220

gataatcacc gcaagacggg caacaggatt caatcttaag aaactttatt gccaaatgtt 5280

tgaacgatcg gggatcatcc gggctctggt cggggaactcc acgaaaatat ccgaacgcag 5340

caagatatcg cgggtgatct cytcttctg tgggcagtcg ccgcogaacc cggtgatgtg 5400

gagccggggc ccgatcatat tgcgctcag gatcgtggcg ttgtgcttgt cggccgttgc 5460

tgctgtaatg atatoggcac ctccgacggc ctgttccgca gagatcccggt gggcgagaa 5520

ctccagcatg agatccccc gctggaggat catccagccg cgtcccgga aaacgattcc 5580

gaagcccaac ctttcataga aggcggcggt ggaatcgaaa tctcgtgatg gcaggttggg 5640

cgctgccttg tcggtcattt cgaaccccag agtcccgctc agaagaactc gtcaagaagg 5700

cgatagaagg cgatgcgctg cgaatcgga ggcgcgatac cgtaaacgac gaggaagcgg 5760

tcagcccatt ccgcgccaag ctcttcagca atatcacggy tagccaaccg tatgtcctga 5820

tagcgttcgg ccacaccag ccggccacag tcgatgaac cagaaaagcg gccattttcc 5880

accatgatat toggcaagca ggcacgcca tgggtcacga ccgatcatc gccgtcgggc 5940

atgcgcgect tgagcctggc gaacagttcg gctggcgcca gccctgatg ctcttcgctc 6000

agatcatcct gatcgacaag accggcttcc atccgagtac gtgctcgctc gatgcgatgt 6060

ttcgcttggt ggtcgaatgg gcaggtagcc ggcataagcg tatgcagcgg ccgcattgca 6120

tcagccatga tggatacttt ctggcgagga gcaaggtgag atgacaggag atcctgcccc 6180

ggcacttcgc ccaatagcag ccagtcctt ccgcttcag tgacaacgct gagcacgct 6240

gcgcaaggaa cgcccgtcgt ggcagccac gatagccgcg ctgcctcgtc ctgcagttca 6300
ttcagggcac cggacaggtc ggtcttgaca aaaagaacgc ggcgccctcg cgtgcagacg 6360
cggaacacgg cggcatcaga gcagccgatt gtctgttggtg ccagtcata gccgaatagc 6420
ctctccacc aagcggccgg agaacctgcg tgcaatccat cttgttcaat catgcgaac 6480
gatccagatc cggtgagatc tatttggatt gagagtgaat atgagactct aattggatc 6540
cgaggggaat ttatggaacg tcagtggagc atttttgaca agaaatattt gctagctgat 6600
agtgaacctta ggcgactttt gaacgcgcga taatgggttc tgaagtatgt gcttagctca 6660
ttaaactcca gaaaccgcgc gctgagtggtc tcttcaacg ttgcggttct gtcagttcca 6720
aacgtaaac ggcctgtccc gcgtcatcgg cgggggtcat aacgtgactc cottaattct 6780
ccgctcatga tcagattgto gtttccgcgc ttcagtttaa actatcagtg tttgacagga 6840
tatattggcg ggtaaaccta agagaaaaga gcgtttatta gaataatcgg atatttaaa 6900
ggcggtgaaa aggtttatcc gttcgtccat ttgtatgtc atgccaacca cagggttccc 6960
cagatctggc gccggccagc gagacgagca agattggcg ccgcccgaaa cgtaccgaca 7020
gcgcgccagc cacaggtgcg caggcaaatt gcaccaacgc ataccgcgc agcagaatgc 7080
catagtggtc ggtgacgtcg ttcagtgaa ccagatcgcg caggaggccc gccagcaccg 7140
gcataatcag gccgatgccg acagcgtcga gccgcagagt gctcagaatt acgatcaggg 7200
gtatgttggt tttcagctct gccctccgga ccagcctcgc ctggtccgat tgaacgcgcg 7260
gattctttat cactgataag ttggtggaca tattatgttt atcagtgata aagtgtcaag 7320
catgacaaag ttgcagccga atacagtgat ccgtgccgc ctggacctgt tgaacgaggt 7380
cggcgtagac ggtctgaca caccgaaact gccggaacgg ttgggggttc agcagccggc 7440
gctttactgg cacttcagga acaagcgggc gctgctcgac gcactggcg aagccatgct 7500

ggcggagaat catacgcatt cggtgccgag agccgacgac gactggcgct cattttctgat 7560
cgggaatgcc cgcagcttca ggcaggcgct gctgcctac cgcgatggcg cgcgcatcca 7620
tgccggcagc cgaccgggag caccgcagat ggaacggcc gacgcgcage ttgccttcc 7680
ctgcgaggcg ggtttttcgg ccggggacgc cgtcaatgag ctgatgacaa tcagctactt 7740
cactgtttgg gccgtgctg agggacggc cggcgacagc gatgccggcg agcgcggcgg 7800
caccgttgaa caggctccgc tctgcgcgt gttcggggcc gcgatagacg ccttcgacga 7860
agccggtccg gacgcagcgt tcgagcayyy actcgggtg attgtcgatg gattggogaa 7920
aaggaggctc gttgtcagga acgttgaagg accgagaaag ggtgacgatt gatcaggacc 7980
gctgcgggag cgcacccac tcactacagc agagccatgt agacaacac cctcccccct 8040
ttccaccgag tcagacgccc gtagcagccc gctacgggct ttttcatgcc ctgccctagc 8100
gtccaagcct cagggccgag ctgcgctct ctggcggcct tctggcgctc ttccgcttcc 8160
tcgtctactg actcgtcgc ctgggtcgtt cggctcgggc gagcggtatc agctcactca 8220
aaggcggtaa tacggttato cacagaatca ggggataacg caggaaagaa catgtgagca 8280
aaagggcagc aaaggccag gaaccgtaaa aaggccgcgt tgctgyuylt tttccatagg 8340
ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa gtcagagggt gcgaaaccg 8400
acaggactat aaagatacca ggcgtttccc cctggaagct cctcgtgag ctctcctgtt 8460
ccgaccctgc cgttaccgg atacctgtcc gcctttctcc cttcggaag cgtggcgctt 8520
ttccgctgca taacctgtct tcggggtcat tatagcgatt ttttcgggat atccatcctt 8580
tttcgcacga tatacaggt tttccaaag ggttcgtgta gactttcctt ggtgtatcca 8640
acggcgctag ccgggcagga taggtgaagt agggccaccc gcgagcgggt gttccttctt 8700
cactgtccct tattcgcaco tggcggtgct caacgggaat cctgctctgc gaggtggcc 8760

ggctaccgdc ggcgtaacag atgagggcaa gcgatggct gatgaaacca agccaaccag 8820

gaagggcagc ccacctatca aggtgtactg cttccagac gaacgaagag cgattgagga 8880

aaaggcggcg gcggccggca tgagcctgtc ggccctacgt ctggccgtcg gccaggcgta 8940

caaatcacg ggcgctgtgg actatgagca cgcccgagag ctggcccgca tcaatggcga 9000

ctggggcgc ctgggcggcc tgctgaaact ctggctcacc gacgaccgcg gcacggcgcg 9060

gttcggtgat gccacgatcc tcgccctgct gggaagatc gaagagaagc aggacgagct 9120

tggaaggtc atgatggggc tggtcgccc gggggcagag ccattgacttt ttttagcgct 9180

aaaacggcg ggggggtgccc gtgattgcca agcagctccc catgcgctcc atcaagaaga 9240

gcgacttcgc gtagctgggt aagtacatca ccgacgagca aggcaagacc gagcgctttt 9300

gcgacgtca ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca 9360

aacgcgccag aaacgcgcgc gaagccgtgt gcgagacac gcggccgcgc gcgttgtgga 9420

tacctcgcg aaaacttgcc cctcactgac agatgagggg cggacgttga cacttgaggg 9480

gccgactcac ccggcgcgcc gttgacagat gggggcagg ctcgatttgc gcggcgacg 9540

tggagctggc cagcctcgca aatggcgaa aacgcctgat ttacgcgag tttccacag 9600

atgatgtgga caagcctggg gataagtgc ctgcggtatt gacacttgag gggcgcgact 9660

actgacagat gggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg 9720

gcgcacctat tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaaggg 9780

ttccgccgt ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt 9840

ataaaccttg tttttaacca gggctgcgcc ctgtgcgctg gaccgcgcac gccgaaggg 9900

ggtgcccccc cttctogaac cctccggcc cgctaacgcg ggcctcccat ccccccagg 9960

gctgcgcccc tcggcgcgga acggcctcac cccaaaaatg gcagcgctgg cagtccttgc 10020

cattgccggg atogggggcag taacgggatg ggcgatcagc cagagcgoga ogcccggaag 10080

cattgacgtg ccgcagggtg tggcatcgac attcagcgac caggtgccgg gcagtggagg 10140

cgccggccctg ggtggcgggc tgcccttcac ttccggccgtc ggggcattca cggacttcat 10200

ggcggggcgg gcaattttta ccttgggcat tcttggcata gtgtcgcggt gtgcgtgct 10260

ogtgttcggg ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg 10320

aggtatgaaa acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa 10380

agulaucag acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac 10440

aatactgata agataatata tctttttat agaatatata gcgctatgta aggatttccg 10500

ggggcaaggc ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa 10560

cttgcatgga ctaatgcttg aaaccaggga caataacctt atagcttgta aattctatca 10620

taattgggta atgactcaa cttattgata gtgttttatg ttacgataat gcccgatgac 10680

tttgtcatgc agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgcag 10740

gtgtgcctc agattcagggt tatgcgctc aattcgctgc gtatatcgct tgcgtattac 10800

gtgcagcttt ccttcaggcg gggattcata cagcggccag ccattccgta tccatatcac 10860

cacgtcaaa ggtgacagca ggctcataag acgcccagc gtccgcatag tgcgttccac 10920

gaatacgtgc goacacaacc tcttcgggag actgtcatc gcgtaaaaca gccacgctg 10980

gcgcgattta gccccagacat agcccccactg ttcgtccatt tccgcgcaga cgatgacgtc 11040

actgcccggc tgtatgcggc aggttaccca ctgcggcctg agttttttaa gtgacgtaaa 11100

atcgtgttga ggccaacgcc cataatcggg gctgttgccc ggcattcaac gccattcatg 11160

gccatatcaa tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt 11220

tgccatgttt tacggcagtg agagcagaga tagcgctgat gtcggcggtt gcttttgcg 11280

ttacgcacca ccccgctcagt agctgaacag gagggacagc tgatagacac agaagccact 11340
gggacacctc aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat 11400
tgtgggtttca aaatcggtc cgtcgatact atgttatatcg ccaactttga aaacaacttt 11460
gaaaagctg ttttctggta ttttaagggtt tagaatgcaa ggaacagtga attggagttc 11520
gtcttgttat aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat 11580
aataaatggc taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct 11640
gcgtaaaaga tacggagga alytctctctg ctaaggtata taagctggtg ggagaaaatg 11700
aaaacctata tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac 11760
gggaaaagga catgatgcta tggctggaag gaaagctgoc tgttccaag gtcctgcact 11820
ttgaacggca tgatggctg agcaatctgc tcatgagtga ggccgatggc gtccctttgct 11880
cggaagagta tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgoa 11940
tcaggctctt tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc 12000
gottagccga attggattac ttactgaata acgatctgyc cgatgtggat tgcgaaaact 12060
gggaagaaga cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa 12120
agccogaaga ggaacttgtc ttttcccaog gcgacctggg agacagcaac atctttgtga 12180
aagatggcaa agtaagtggc tttattgatc ttgggagaag cggcaggggc gacaagtggc 12240
atgacattgc cttctcgctc cggtcgatca gggaggatat cggggaagaa cagtatgtcg 12300
agctattttt tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatatatt 12360
tactggatga attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag 12420
cgcaccgact tcttcgcgat caagtgtttt ggctctcagg ccgagggcca cggcaagtat 12480
ttgggcaagg ggtcgctggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag 12540

gacggccaga cgggtetacgg gaccgacttc attgcogata aggtggatta tctggacacc 12600
aaggcaccag gcgggtcaaa tcaggaataa gggcacattg ccccgcgctg agtcggggca 12660
atcccgaag gagggatgaat gaatcggacg tttgacogga aggcatacag gcaagaactg 12720
atcgacgcgg ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatcgct 12780
gcgccccgcg aaaccttcca gtccgtcggc togatggctc agcaagctac ggccaagatc 12840
gagcgcgaca gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag 12900
cgttcgcgtc gtctcgaaca ggaggcggca ggtllggcga agtcgatgac catcgacaacg 12960
cgaggaacta tgacgaccaa gaagcgaaaa accgcggcg aggcacctggc aaaacaggtc 13020
agcgaggcca agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaatgcag 13080
ctttccttgt togatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg 13140
gccgcctctg cctgtttcac caccgcgaac aagaaaaatcc cgcgcgagggc gctgcaaaac 13200
aaggtcatit tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg 13260
gccgacgatg acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc 13320
ggcgagccga tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat 13380
ggccggtatt acacgaagcg cgaggaaatgc ctgtcgcgcc tacaggcgac ggcgatgggc 13440
ttcacgtccg accgcgttgg gcaacctggaa tcggtgtcgc tgcgtcaccc ctcccgctc 13500
ctggaccctg gcaagaaaaa gtcccgttgc caggtcctga tocgacgagga aatcgtcgtg 13560
ctgtttgctg gcgaccacta caccgaaatc atatgggaga agtaccgcaa gctgtcgcg 13620
acggcccgcg ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg 13680
gaaaccttcc gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag 13740
gtcggcggaag cctgcgaaga gttgcgaggg agcggccctg tgggaacacg ctgggtcaat 13800

gatgacctgg tgcattgcaa acgctagggc cttgtggggc cagttccggc tgggggttca 13860
gcagccagcg ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc 13920
gctcagtatc gctcgggacg cagggcgcg cttacgaabt gccgataaac agaggattaa 13980
aattgacaat tgtgattaag gctcagattc gacggcttgg agcgcccgac gtgcaggatt 14040
tcgcgagat ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg 14100
agcacgagga gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat 14160
tcggcgccca catcgacggc gagatcattg ggtgttcggt cttcaaacag gaggacggcc 14220
ccaaggacgc tcacaaggcg catctgtccg gogttttcgt ggagcccgaa cagcgaggcc 14280
gaggggtcgc cggtatgtcg ctgcgggcgt tgcggcgggg tttattgtc gtgatgatcg 14340
tcgcacagat tccaacggga atctggtgga tgcgcattt catctcggc gcacttaata 14400
tttcgtatt ctggagcttg ttgtttattt cggctctacc cctgcggggc ggggtcgcgg 14460
cgacggtagg cgctgtgcag ccgctgatgg tcgtgttcac ctctgccct ctgctaggta 14520
gcccatacgc attgatggcg gtctctgggg ctatttcggg aactgcgggc gtggcgctgt 14580
tggtgttgac accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg 14640
cggtttccat ggcgttcgga accgtgctga cccgcaagt gcaacctcc gtgcctctgc 14700
tcacctttac cgcctggcaa ctggcgccg gaggacttct gctcgttcca gttagctttag 14760
tgtttgatcc gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggtcgc 14820
gcctgatcgg agcgggttta acctacttcc tttgggtccg ggggatctcg cgactcgaac 14880
ctacagttgt ttccctactg ggctttctca gccccagatc tggggtcgat cagccgggga 14940
tgcatcaggc cgacagtcgg aacttcgggt ccccgacctg taaccattcg tgagcaatgg 15000
ataggggagt tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcttc 15060

agcggccttta tccagcgatt tcttattatg tcggcatagt totcaagatc gacagcctgt 15120
 cacggtaag cgagaaatga ataagaagcg tgataattcg gatctctgcy agggagatga 15180
 tatttgatca caggcagcaa cgctctgtca tcgttacaat caacatgcta cctccgcga 15240
 gatcatccgt gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac 15300
 atgagcaaaag totgcccgcct tacaacggct ctcccgctga cgcgcgtccg gactgatggg 15360
 ctgctgttat cgagtgggtga ttttgtgccg agctgcgggt cggggagctg ttggctggct 15420
 ggtggcagga tatattgtgg tgtaaacaaa ttgacyctta gacaacttaa taaccacattg 15480
 cggacgtttt taatgtactg ggggtgtttt tottttcacc agtgagacgg gcaacagctg 15540
 attgccttc accgcctggc cctgagagag ttgcagcaag cgytccaagc tggtttgccc 15600
 cagcaggcga aaatcctggt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca 15660
 aaagaatagc ccgagatagg gttagtgtt gttccagttt ggaacaagag tccactatta 15720
 aagaacgtgg actccaagct caaaggcgga aaaacgctct atcaggcgga tggcccacta 15780
 cgtgaacat caccocaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg 15840
 aaccctaag ggagccccg atttagagct tgacggggaa agccggcgaa cgluyuyaga 15900
 aaggaaagga agaaagcgaa aggagcgggc gcaattcagg ctgcgcaact gttgggaagg 15960
 gcgatcgggt cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag 16020
 gcgattaagt tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag 16080
 tgaattcgag ctcggtaccc ggg 16103

<210> 63

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 63

ggcgtacttg aaggaaccct taccg

25

<210> 64

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 64

attgatgctc cgggtcacgc tgatt

25

<210> 65

<211> 500

<212> DNA

<213> Blakeslea trispora

<400> 65

aatctatata atgtccata gactcacatt gatattgtcg aagatttcga tgcgtactta 60

gtagagcaac tacaaaagt agcagagaag catgatttct taatctttga agaccgcaag 120

tttcgagata tcggtatgtg aattctatct atttttttct tgatgtgtgc atggatgact 180

catgatcata ttcttaggta atactgtcaa gcatcaatat ggcaaggcgc ttacaagat 240

tgcttcttgg tctcatatta ctaatgtcga cacagttcct ggagaaggta ttatcaaggg 300

acttgccgaa gloggcctcc ctcttggtcg tggttgcct ttgctagcag aaatqtcac 360

tcaaggtgca ttaactaagg gtatttacac tgcgaatct gtcaatatgg ctgcgcgcaa 420

caaagatttc gtttttggct ttattgcaca acacaaaatg aatcagtatg atgatgagga 480

ttttgttgtc atgtcgctg

500

<210> 66

<211> 611

<212> DNA

<213> Blakeslea trispora

<400> 66

gagattaaaa tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt 60

otttttataa ataatgctgt gactttcttt ttcoaatgat atgatatatg catatgatag 120

atatacaagc aatcttcttc aaggagtttg aaattttgtc ctocaggagc aaaaaaagt 180

ttttttttat aatgttttgt aacaagaat agttaccaat ttgcttttgt cttacgtgt 240

gcaagtttat atcgttttca atttctttgt cttacattt tctttgtctt ttatctttoc 300

tcatttagtc ttggggagaa ttaggaaaag ggagcggaaa ggtaagaaat gottgcgtat 360

tttactaatt oggcaaacat ccaattttgc aaacagcagc ctgtgcaacg ctctcgagat 420

gacagtatct ttgattacac totaaatctc gatgaccoga caaaaagag gaaacaaga 480

aataatcttg tgcattogaa tatgatgaa gatttttttc ccttattctt aaatgttgac 540

atagcgtgta tgttatataa aaaaaagaa attgtacaaa cttctttttc ttctcttttt 600

attttatctc t 611

<210> 67

<211> 720

<212> DNA

<213> Blakeslea trispora

<400> 67

atgtcaatc tcaattatct ggaatttcac ctctactata cactacctgt cttggggca 60

ttgtgtttgc tgcataagcc gtttcaactc cagcaagaca atctcaagta taaattttta 120

atgttgatgg cgcctctac cgcctcgatt tgggacaatt atactgttta tcatcgcgct 180
 tgggtgtact gtctactctg tgttggtgct gtcattggct atgtacctct agaagaatac 240
 atgttcttta tcatcatgac tttaatgact gtgcggttct caaactttgt tatgcgttgg 300
 cacttgcata ctttctttat tagaccaac acttcttga agcaaacact attagtacgc 360
 cttctgcctg tttaacgttt attggcaatc acttatcatg cttggcaact gacactgcca 420
 aataaacctt cattttatgg ttcatgcac ctttgggtatg cttgtcctgt gttggctatt 480
 ctttggctgg gtgtggcga atatatcttg cgtcgacctg tggctgtcct tttgtctatt 540
 gttatcccta gtgtatacct atgttgggct gatatacttg ctattagtgc tggcacatgg 600
 catatttctc ttagaacaag cactggcaaa atggtagtac ccgatttacc tgtagaagaa 660
 tgccgtgttt ttactttgat caacacagtc ttggtttttg ctacctgtgc tatagaccgc 720

<210> 68

<211> 1089

<212> DNA

<213> *Blakeslea trispora*

<400> 68

ctgtacaaat catctgttca aaatcaaac cctaacaag ccatttcctt ttccagcat 60
 gtcaaagagc tagcatgggc cttctgtctt cctgacaaa tgcacaaca tgaattgttt 120
 gatgatctta ctactagctg ggatatttta cgtaaagcct caaagtcatt ctatactgca 180
 tctgcgcttt ttocaagtta tgtaogtcaa gaottgggtg ttctctatgc tttctgcaga 240
 gctaccgatg acctgtgcga tgatgaatoc aaatctgttc aagaagaag agaccaatta 300
 gatcttactc gacaatttgt tegtgtctc tttagccaaa agaccagtgc gcctattgtg 360
 attgattggg aattgtatca aaaccaactt cctgcttctt gtatatcagc ctttagagcc 420
 ttactcgcgc ttogccatgt ccttgaagta gacctgtag aagaactatt agatggttac 480

aaatgggatac ttgagcgctcg tcctatcctt gatgaacaag acttggaggc atactctgct 540
 tgtgtggcca gtagtgtggg tgaaatgtgc acacgtgtga ttcttgtctca agacacaaag 600
 gaaaatgatg cttggataat tgaccgtgca cgtgagatgg ggctgggtgct acaatacgtt 660
 aacattgctc gagacattgt gactgatagc gagactctgg gtctgatgta tctgcctcaa 720
 caatggctta gaaaagaaga aacagacaaa atacagcaag gcaacgcccg tagcctaggt 780
 gatcaagac tgttgggctt gtctctgaag cttgtaggaa aggcagacgc tatcatgggtg 840
 agagctaaga agggcattga caagttgccg gcaaacctgc aaggcgggtgt acgagctgct 900
 tgccaagtat atgtgcgaat tggatctgta ctcaagcagc agaagacaac atactctaca 960
 agagctctac taaaaggaag cgaacgtgcc aagattgctc tgttgagtgt atacaacctc 1020
 tatcaatctg aagacaagcc tgtggctctc cgtcaagcta gaaagattaa gagttttttt 1080
 gttgattag 1089

<210> 69

<211> 611

<212> DNA

<213> *Blakeslea trispora*

<400> 69

agagataaaa taaaagaga agaaaagaa gtttgtacaa ttcttttttg tttatataac 60
 atacaagcta tgtcaacatt tagaataagg gggaaaaaat cttccatcat attogaatgc 120
 acaagattat ttctttgttc gctotttttg gtctgggtcat cgagatttag agtgtaatca 180
 aagatctgt catctcgaga gogttgcaca ggctgctgtt tgccaaattg gatgtttgac 240
 gaattagtaa aatacgcaag cattttctac ctttcgctc ccttttctca attctcccaa 300
 agactaaatg aggaagata aaggacaaag aaaatgtaaa gacaaagaaa ttgaaaacga 360

tataaaacttg cagcacgtaa gaccaaagca aattggtaac tattcttgtg tacaaacatg 420
 tataaaaaaa aacttttttt tgctcctgga ggacaaaatt tcaactcct tgaagaagat 480
 tgcttgata tctatcatat gcatatatca tatcgatgga aaaagaaagt caggcatgta 540
 tttataaaaa gaagaatgtg ccatgcttcc gaatttcttt tcactttctt ttccttatct 600
 attttaatat c 611

<210> 70

<211> 882

<212> DNA

<213> *Haematococcus pluvialis*

<400> 70

atgctgtcga agctgcagtc aatcagcgtc aaggcccgcc gcgttgaact agcccgcgac 60
 atcacggcgc ccaaagtcgt cctgcattgt cagcgggtgt cgttagtttg gctgcgagtg 120
 gcagcaccac agacagagga ggcgctggga accgtgcagg ctgcgcgcgc gggcgatgag 180
 cacagcgccg atgtagcaat ccagcagctt gaccgggcta tcgcagagcg tcgtgcccg 240
 cgcaaacggg agcagctgtc ataccaggct gccgccattg cagcatcaat tggcgtgtca 300
 ggcattgcga tcttcgccac ctactcgaga ttgccatgc acatgacctt gggcgccgca 360
 gtgccatggg gtgaagtggc tggcaactct ctcttggtgg ttggtggcgc gctcggcatg 420
 gagatgtatg ccgcgtatgc acacaaagcc atctggcatg agtcgcctct gggctggctg 480
 ctgcacaaga gccaccacac acctcgcaat ggaccctttg aagccaacga cttgtttgca 540
 atcatcaatg gactgcgcgc catgctcctg tgtacctttg gcttctgggt gcccaacgtc 600
 ctggggggcg cctgcttttg agcggggctg ggcatacgc tatacggcat ggcataatat 660
 tttgtacacg atggcctggg gcacagcgc tttcccadcg ggcccatcgc tggcctgcc 720
 tacatgaagc gcctgacagt ggcccaccag ctacaccaca gcggcaagta cggtgccgcy 780

ccctggggta tgttcttggg tccacaggag ctgcagcaca ttccaggtgc ggcggaggag 840
gtggagcgac tggctctgga actggactgg tccaagcggg ag 882

<210> 71
<211> 528
<212> DNA
<213> *Erwinia uredovora*

<400> 71
atgttgtgga ttggaatgc cctgatcgtt ttogttaccg tgattggcat ggaagtgatt 60
gctgcactgg cacacaaata catcatgcac ggctgggggt ggggatggca tctttcacat 120
catgaaccgc gtaaagggtc gtttgaagtt aacgatcttt atgccgtggt ttttgctgca 180
ttatcgatcc tgetgattta totgggcagt acaggaatgt ggcgcgtcca gtggattggc 240
gcaggatgta cggcgtatgg attactctat tttatggtgc acgacgggct ggtgcacaa 300
cgttggccat tccgctatat tccacgcaag ggctacctca aacgggtgta tatggcgac 360
cgtatgcac acgcgcgcag gggcaaagaa ggttgtgttt cttttggctt cctctatcgc 420
ccgccctcgt caaaacttca ggcgacgctc cgggaaagac atggcgctag agcggggcgt 480
gccagagatg cgcaggcgcg ggaggatgag cccgcacccg ggaagtaa 528

<210> 72
<211> 762
<212> DNA
<213> *Nostoc sp. PCC73102*

<400> 72
atgatccagt tagaacaacc actcagtcac caagcaaac tgactccagt actgagaagt 60
aaatctcagt ttaaggggct ttctattgct attgtcattg ttagcgcagt ggtoattago 120
ctgagtttat taectttocct tgacatctca aagctaaaat ttgggatggt attgcctggt 180

atactatggc aaacattttt atatacgga ttatttatta catctcatga tggcatgcac 240
 ggcgtagtat ttcccaaaa caccaagatt aatcatttga ttggaacatt gaccctatcc 300
 ctttatggtc ttttaccata tcaaaaacta ttgaaaaaac attgggttaca ccaccacaat 360
 ccagcaagct caatagacc ggattttcac aatggtaaacc accaaagttt ctttgcttgg 420
 tatttttcatt ttatgaaagg ttactggagt tgggggcaaa taattgcgtt gactattatt 480
 tataaotttg ctaaatatcat actccatato ccaagtata atctaetta cttttgggtg 540
 ctacccctgc ttttaagttc attacaatta ttctattttg gtactttttt accccatagt 600
 gaaccaatag ggggttatgt tcagcctcat tgtgccaaa caattagcgg tctattttgg 660
 tggtcattta tcacgtgcta tcattttggc taccacgagg aacatcacga atactcctcat 720
 atttcttggg gccagttacc agaaatttac aaagcaaat ga 762

<210> 73

<211> 617

<212> DNA

<213> Haematococcus pluvialis

<400> 73

taggggtcgg aaccaggcac gctggtttca cacctcatgc ctgtgataag gtgtggctag 60
 agcagtcggt gtgagacggg tatgtcacgg togactggto tgatggccaa tggcatcggc 120
 catgtctggg catcacgggc tggttgcctg ggtgaagggtg atgcacatca tcatgtcggg 180
 ttggaggggc tggcacagtg tgggctgaac tggagcagtt gtccaggctg gcgttgaato 240
 agtgagggtt tgtgattgga ggttgtgaag caatgaatcc gcccatatto tatttgtggg 300
 agctgagatg atggcatgct tgggatgtgc atggatcatg gtagtgcaga aaactatatt 360
 cacctagggc tgttggtagg atcaggtgag gccttgaca ttgcatgatg tactogtcat 420

gggtgtgttg tgagaggatg gatgtggatg gatgtgtatt ctacagacgta gaccttgact 480
 ggaggcttga tcgagagagt gggccgtatt ctttgagagg ggaggctcgt gccagaaatg 540
 gtgagtggat gactgtgacg ctgtacattg caggcaggtg agatgcactg tctcgattgt 600
 aaaatacatt cagatgc 617

<210> 74

<211> 1208

<212> DNA

<213> *Haematococcus pluvialis*

<400> 74

attgtgactg atagcgagac tctgggtcga tgttatctgc ctcaacaatg gcttagaaaa 60
 gaagaacag aacaaatata gcaaggcaac gcccgtagcc taggtgatca aagactgttg 120
 ggcttgtctc tgaagcttgt aggaaggca gacgctatca tggtagagac taagaagggc 180
 attgacaagt tggcggcaaa ctgtcaaggc ggtgtacgag ctgcttgcca agtatatgct 240
 gcaattggat ctgtactcaa gcagcagaag acaacatata ctacaagagc tcactataaa 300
 ggaagcgaac gtgccaagat tgctctgttg agtgtataca acctctatca atctgaagac 360
 aagcctgtgg ctctccgtca agctagaaag attaagagtt tttttgttga ttagtgaatt 420
 ttgtttttat ttatgtctga tagttcaata aagagacaac acatacaata taaatcatt 480
 gtctttaaat gttaatttag tagagtgtaa agcctgcatt tttttgtac gcataaaca 540
 tgaattcacc ccgcttctgg tttttaata attatgtcaa actagggaaa attctttttt 600
 ttctcttcgt tctttttttg gcttgttgtg gagtacagg cttgtcttca gattgataga 660
 ggttgtatac actcaacaga gcaatcttgg cagcttcgct tctttttaga tgagctcttg 720
 taggatattg tgtctctcgc tgcttgagta cagatccaat tgcagcatat acttggcaag 780

cagctcgtac accgccttga cagtttgcog gcaacttgtc aatgccotte ttagctctca 840
ccatgatagc gtctgccttt cctacaagct tcagagacaa gcccaacagt ctttgcacac 900
ctaggctacg ggcgttgccct tgctgtattt gttctgttct ttcttttcta agccattggt 960
gaggcagata acatgcagccc aacatcctcg agccatacta cagcataaaa ggatacgttt 1020
tctttaacag caattttacc ttttgtatc agcacatata aaaaaaaga aatttaagat 1080
gagtaggact tccattctct caaaaatttt attcaatcca taaatgaatt atttttggac 1140
aaaaaagaaa gattatgcct gattttctct attttttttt tttttacaac tccaccaata 1200
ctttctag 1208

<210> 75

<211> 6316

<212> DNA

<213> *Blakeslea trispora*

<220>

<221> misc_feature

<222> (2694)..(2694)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (4263)..(4263)

<223> n is a, c, g, or t

<400> 75

aaggatgaag aatccaactc taataaaaat cttatggata tctttgatcg actcaaaaag 60
gctttcaatg ctattgctat taaaaaaaaa gagagagaga gaaotatgag caaaaggact 120
ctatgccaaag atggcaaaaa ggcaccagaa acccttagtt tattattgca taatccagtc 180
gagctagtac ttctgtagct caagcttaac cgaggatcct ggaatcaact cgtctcgtca 240

ctcttgcga tgatcctaga aatggtatct atggatgta tactaacatt gttatcttct 300

aaggcctcga agatgttatt gttgcggtga taaataggct gctatgtact gaagttgctc 360

tgtaaaatga atctagttca ctgcctactc agcaaatggg tgtttctaata gtctttaaag 420

aaagaaaaaa agatacatat agactaccct tcttttcaag actgtaatcg agaatcggcc 480

gattggtttat tacaattaga cgtcgggaat aagcaaaagg attcatcttt gtaataaaga 540

gactggtgca tatgaaagca aggatcgat caaggaatag ttttgatcga gcataccag 600

caaatgctgc taatgtttgc ttcttctttg ctctctgaga ttgaatggga tgtgcctaga 660

gcattgctat ttttaagtg atactttaga ttgtgtctt tagatttggt tcattttatt 720

tagtcaagaa agatcccoct ttctctatgt atgctaagaa gaaggagcaa gaagtgtatt 780

tacaagttgg aatgagattg aaatattgta cataataata ataaaaagaa aggtagatca 840

aaaaaatgt tctgcctatt gtaagaaatc gggaccaaca ggtgcttgat aaccagaagt 900

agcttccaat tcaggtagag gctctagga caaatacaca attatgacag gaattttctt 960

gttgacttga acactacaag agaaacgggt cagcacaaaa tccgaaaaaa aaagaaacg 1020

gaccattcat gtcttaccta tctagctctt tgtcttcaat tgcattccat tgcacaacca 1080

cagatcgct tcccaattga gtatattgat gaagtgttcc ctgcattttt cgcttgacta 1140

attccactac agtcacagtc ttattaatgt tttgtcctt accagtcagg ataatatgat 1200

ctttttgctt ctctctacaa aaaaataatt cttgttttga ataaaaaaa caaatattta 1260

aagaaactac ttgatgacg gtacctggaa taactcgaga cacacatcta catatcggtt 1320

gattttattg tggotaatto gaaocctatt ttctgctggt ggggctgtt gactttcagt 1380

tgctgagagc tcttctttgc ttcttttata gtcttcact atgattttaa tcaagaaggt 1440

aagtcagtga tgattgttac aagctatata tcttgaaaaa gaacagagag gtattattat 1500

cagatgcaac atggttttct gtatcatttt catttcagtt tctctgttca aaaaaaaaa 1560

gaacactttc tctttccact cctcaaattt tttctgctaa actcctcgca aaacatgtat 1620

ttgctttaaa ctacaagttg caattgtctg atttagcaat ttcaatagc cttttgtgaa 1680

tccacccaaa aataaacaag tgcttgagta tacttggtt cagttcaaaa gaaagcaagc 1740

tttttttttt ctttctggg aaagaaaaaa aaatatgtt gagccatcct ttaccagcag 1800

tatgcgagct acgacatagc tggcttaaca atgactgcaa gcaatagatc gagcttagtc 1860

ttctattgc ttcyttgttt gatctatgtt cggccttacg ctgacctatc caatactcga 1920

gataggcaac aagatttcga acagtaatga aataaatttc ggataacagt tgtggatgag 1980

gaagagaaag cgacttgaac tcgagaaact ttgttgaat gaaatccgac cttttacgtg 2040

atcatcatgt attatcctct ttttcttttt ttctgtagt aattacttac tgattgcgct 2100

caagtgcgct ctttataaag aagaaaaaaa aatattagaa ctttcaaaaa atataactga 2160

aaataaaagt gtggctcgga gagcaaatc cacatccttt gtcttcgctt tggtaacacg 2220

gttaataagc cactataggt gaataatgat catttctgag aataaagcgc ggcttgaagc 2280

ttatatccat atcaggatc atattaggca caactcaca ttgaggttcc agaagtcca 2340

attttttttt cctgatagcc tgtccaatta agatcaaaaa ccactgagtt ttctctatat 2400

attttttttt ttacataattc ttaactcttc ttctctctc ttctctctc ttctctttttg 2460

gcttgcaaaa aaaatcttta gtaataccaa agaaagcaaa ccttttcctt ttcttatttc 2520

cttgcttggt ttttaatttt tgatttctct atgctttaa taccatttc ttctctctt 2580

ctgctattac ctatcttttc attctctctc cccctctctc tcttggtcta taaacatcat 2640

gaagtcctct tttaaaagtt cgcttgacat ttatgtgtt tatatacagc atctgtgtt 2700

ttccaagtg ttcattcttg cttttgttct ttcgatttc ctcaacactt atctactgaa 2760

cgcttcgaag caacagccca aagtataat caaaaagggtt attgagcggg tagaagtacc 2820

aagtagagaa caacctaaat cagtcataaa gccctcctcc aagaacact cttctcatca 2880

tcagtcgat gtcattcgcc ctcttgatga agtattgggt ttgctcgaa caccogaggc 2940

cttgactgat gaagagatca tctctattgt tcaagctggg aaaaaggccc cctatgctct 3000

tgaaaaggtc ttgggcgatt tagagcgcgc tgtccatato cgtcgtgctt tgatctcccg 3060

tgactctcgt acgaaaactt tggaagacag tatgcttccc gtgaaaact atcattatga 3120

taagtcatg ggtgcttgtt gtgaaaagt cattgggtat atgcctattc cagtaggtgt 3180

cgcggttaag aagttcaaca agtcgcgata ttgacaagt tgctcatcat ttctgaaaca 3240

ggctccttgg tgattgatgg tgattctatt catattccca tggcaactac ggaaggttgt 3300

ttagttgctt ctactgccag aggttgtaaa gcaatcaatg ctggtggtgg tgccaacaca 3360

attgttggtg ctgatggatg gactcgagggt ccttggtcgc aatttctac aatcactcgc 3420

gctgctgact gtaaacgatg gattgaacaa gagggggaag ctatcgtgac cgaagcattc 3480

aattcaactt ctctgtttgc tcgtgttcgt aaattgaaag ttgctcttgc cggctgtcta 3540

gtctacatcc gtttctctac cactacaggt gatgcaatgg gcataacat gatctccaag 3600

ggttggtgaa aggttttaag caagattgct gagagataac ctgatatgca gatcatttct 3660

ctttctggta actattgtac tgacaagaaa cctgctgcta tcaactggat tgaaggacgt 3720

ggtaaatctg ttgttgctga sgctgtcatc cctggtaggg ttgtcgaaaa ggtattgaag 3780

acctctgtta gtgcttttgg tgagctgaac atctctaaaa acctgggtgg ttctgctatg 3840

gctggctcgg tcggtggctt taaogctcat gctgctaata ttctaactgc catttacctt 3900

gctactggtc aagatctcgc tcaaaaatga sagagttcta actgtattac ttgatgaaa 3960

gctgtcaatg gcgaaagaga ccttcataac tottgtaaca tgccctgtat tgaagtaggc 4020

accattgggtg gtggtactat ttgcctect caacaagcca tgttggaatt cattgggtg 4080

cggtgtcctc accctaccga acctgggtgc aatgccgwc gccttgctcg tgttatctgt 4140

gcctctgtga tggctggtga attgtcttta tgtgcagctt tggctgctgg tcactctgta 4200

aaggcacaca tggctcataa tcgtaatacc actgctgctg ccgctgttgt tctgtccct 4260

aanggcatag ttgatgtctc tacacotcct gctacacctg cagaaaagaa tgatcctatt 4320

cctggaagtt gtatcaagtc atagaattaa tattatatat atatcatata caaaaaaag 4380

aaaaaaaa cactacatct atttattlll clccatgtac acacacacac acacataaa 4440

aaactcttta ttttccaata ttttgctttt ataaaatao ttatttcatt ctaataaaac 4500

tgtttttttt tattaatcat caaacctgc tgagagctgt gcaatatcat ctatgttttc 4560

atggtttaac tctggtatcg gwcgagcctc ctctgtactt gaagtttgta ggcagttttt 4620

atttaaggct gctggctgat catgatcatc akcaaacctg acagcatgaa gttttgactg 4680

atgagcaatt tcactaaggc cagaatctga actctttgc ttctactat tgaccatatt 4740

gtcttaggtt ggaatgagtg aatagcgtct tgtcatatgt aacacagaat caacaatato 4800

ctggtgatga aactcgcca aacatagcgc cttttccccc caacaattat aataatcaaa 4860

atgagaatga catgtacggt tttctcgtat gacaatatcc aacgtcttgt cataatctc 4920

tgtgcyata ccattcatct tttggaagaa cgcacggtag ctctcacaag ctgtcctcag 4980

agagtccgtt gccatgtttc ccaatgctcc tggcaagtcg aaatgaagtt gtcgaatctg 5040

gcgatgatg tctacaatgt cgcctgtttc tttoattaga tcaagcattc gtgtagccca 5100

aatgatgtct atgttatgat tttctttcat tccagtaata actatagttt ctgcgcaaat 5160

cgaatgastg atggagtaaa ttcataaaa gtgcaagtaa tacatacagt gcttgaagaa 5220

atcttggtga gcacgcctat attatgtaat ataggatoga ttctcgaaac tcgacataac 5280

caccaggctt tagcaagcgt tttatttcat tcatgacaag ctattgttaa ttcytgctta 5340
ataaaacaaa atgaaaaaaa cataccccc tcmaaaccta cttccoactc ttgattggaa 5400
aaacaggat agacgtgacg catatgtata taatcaaac actcatcagg ataggggtaaa 5460
ccattgagca catcgcattg ggtgaagaaa gtattaggag gcttgatggc tgtaggatat 5520
ataggtgcaa tatcaatacc gtaaaactca gcatttggga attctgtagc catctccaga 5580
atccaagtac ctgtgcaca agcaacatca agcactttag gtaaggggat acattgttgt 5640
tcttlytlyt gttgttgaca atcacttgag tctgagttc gtttggattg ttttaagtgc 5700
aataattctt ttacagggtc tgagaaatta cgtcaataa gatacttgta aataaaatgc 5760
taaaaataaa aacaatagaa aaaaaattg acgtctcatt cattactatg gaaataactg 5820
caaaatctta ccaactgtac aagtctatct tgctcaatct catcgtttgg cagaatgtat 5880
ttattgttgt agtattcata tcttctacca ttcattgat aactgtcgct tctaagtctc 5940
tgagggtgaag tacttgtagg tgaaggtgga agtgacgcaa ttttgtcaag cttaacagga 6000
tcctctcgcc tacatgtttt ctgcataatca ggaaaatctt gtttatttga aacatcaaca 6060
gtagatgtgg tgtgatcttt tttgaaaata tcgatgcctt cctttgaaag ccttttgaaa 6120
ggctctttta acttttttga gtgagagcta cccatgatag cttatgaaga attaaaaaga 6180
aaaaagcaaa aaaaattaaa aaaaaaaa gtagcaaaaa attctgtcgt aattatacaa 6240
gccaatcaaa atcgaaattc atgcaaggca tagatgttca cgtggatttg atgggtgatc 6300
cttttttttt gcaaga 6316

<210> 76

<211> 1170

<212> DNA

<213> *Thermus thermophilus*

<400> 76

atgaagcgcc ttctcctgag ggaggcctgg ccctacctga aagaacctca gcaagatccc 60

ctegcgctcc tgctggcgtg gggccgggccc caccoccggc tcttccttec cctgccccgc 120

ttccccctgg ccctgatctt tgaccccgag ggggtggagg gggcgctcct cggcgagggg 180

accaccaagg ccaccttcca gtaccgggccc ctctcccgcc tcacggggag gggcctcctc 240

accgactggg gggaaagctg gaaggaggcg cgcaaggccc tcaaagaccc ctctctgccg 300

aagaaagtcc gcggtctacc ggaaggccatg gagggaggcg cccgggcctt ctctggggag 360

tgccgggggg aggagcggga cctggaccac gagatgctcg ccctctcctc gcgcctcctc 420

ggggggggccc tcttcgggaa gccctctec ccaagcctcg cggagcacgc ccttaaggcc 480

ctggaccgga tcatggccca gaccaggagc cccctggccc tcttggacct gggcgccgaa 540

gcccgcttcc ggaaggaccg gggggccctc taccgcgagg cggaagccct catcgctcac 600

ccgccctct cccaccttcc ccgagagcgc gccctgagcg aggcctgac cctcctggtg 660

gggggccaag agacgggtgc gagcgccctc acctggtcct ttctctcctc ctcccaaccg 720

ccggaactggc agaagcgggt ggcgagagc ggggaggcgg ccctcgccgc ctctcaggag 780

ggcctgaggc tctacccccc cgcctggatc ctaccccgga ggctggaaag gccctcctc 840

ctgggagagg accggtctcc ccggggcacc acctgggtcc tctcccccta cytgaccag 900

aggctccact tcccgatgg ggaggccttc cggcccagac gcttctctga ggaagggggg 960

accctctcgg ggcgtactt cccctttggc ctggggcaga ggctctgcct ggggcgggac 1020

ttcgccctcc tcgaggggcc catcgctctc agggccttct tccgcgcctt ccgcctagac 1080

ccctcctcct tccccgggt cctcgcccag gtcacctga ggcaccgaagg cgggcttccc 1140

gcgcggccta gggaggagggt gggggcgtga 1170

<210> 77

<211> 2981

<212> DNA

<213> *Blakeslea trispora*

<400> 77

tctagaattc attccattcg aaaggatcaa cataaccaat ttaatgacta ctagotaatg 60

gatacaata tacgcacaaa aaaagaaaga attctatgat caaagagaac acagacacag 120

agtgatacat ttaaatggtt aagttcttat gatgttaaaa tggtaacttt attattgaat 180

taaatgcgaa tatcgttgct gctttgtact tggaaaacgt taggtaaaag ttggttaatg 240

aaagaagcag gagttgtagt atcatctctt gggaagaaat agaaaaagag gaaagtaaca 300

aagtaacaag caagacaata atagatccaa tggctttcgg tcttacgagt ttgttcagga 360

gcatacttct ttgggtatc ttgtaacttt ctgggtaagg gattctggcc aaagctttta 420

cagacttggt oggaagtaag ctacttcca gcaagaacga taggaacacc agtacctgga 480

tgtgtactac aaagaaaaga gaaatgagta cgtgcgttat taaaaaaaag aaaaaagag 540

ggcaaaagta ttacctagct ccgacaaga aaagattatc ataacggttt gtggaatcct 600

tggtactagg tctgaaccag agaacttgga acacatcatg agaaagacca agaatagaac 660

ctctccaaag gtaaaacttg ctttgccaaa cactaggatc attcacttct tcatgttoaa 720

toaaattagc aaagttgttt actcccaaac gacgttcgat aacttcaga acaatcttgc 780

gtgcacggtt taccaactca ggataatctt ctacagcact gtttcctgtc ttactcttca 840

tatggccaat tggaaaccaac acaataatgg agtccttggt gggaggtgag gcagattcat 900

caattcgaga tggaaacgtt acatagaatg aagcttcaga gggcaaacgg aagtcgttga 960

aaatctcatc aaaactttcc ttgtagggtt cagccaagaa gatattgtgt acgtctaatt 1020

gaggcacett tgttgacatg gaccaataaa acgaaataga tgatgaagtg agtttctttg 1080
aggctaattgt ottctttgtc caattgcaag gaggtaacag atggtgataa gcataaacaa 1140
gatccgcatt acatacgact gcacggcgtt caatgacttc tccgctttcc aaagtgcac 1200
cggttacacg cttgtcttta tcgacagtgt taatttttagc aacaggcgat tgatatctga 1260
attcagcacc gtactttttg gaggcgatag actcaagctt ctgaacaacc atgttgaaac 1320
caccacgagg ataccagata ccttcagcaa actcgggtgta ttgtaacaaa ctgtaaacctg 1380
ctggagcacc ataaggcgac atactatatt ccaaaaatag aaaaatagaac aatgaatctc 1440
aaaattcctt tcacttgcgc tttttcacat ttctcttttc ccaccccgga cgggtctcac 1500
tcattttttt ttcacccac accacgcgtt gtatgtgtac ttaccccata tacattgttt 1560
gaaaagtaaa agccatacgc attttcttgg ttgggaaata ttactggct cgggtcataga 1620
tcttaccaaa caagtgcagc cgaagatttt caggcacata ctgaagacga atcaaatccc 1680
aaatggtttc aaagtgcgc ttgatagcaa taaatgtacc ttgttcataa tggacatgtg 1740
ttctctcat gaaatocaag aatctaccaa atccaagggg accctcaata cgggtocaatt 1800
cgcccttcat cttggttaaa tcggaagaga gttgtacggc atcacogtcg tcaaaatgaa 1860
ccttatagtt attgtcacag cgaagcaaat ccaaatgac accaatacgt tcactcaaat 1920
cagcaaatgc atcttcaaaa agcttaggca tcaaatagag tgagggaccc tgatcaaaagc 1980
gatgaccac gtgatgaatg aatgaacaac ggccacggga aaagtcgttc ttttcaacaa 2040
cagtaactcg aaaaccttca cgagcaagac gagcagcagt agcagttccg ccaataccgg 2100
caccaatgac aacaatatgc ttcttttgat cagacatgag attaaaatag ataaggaaaa 2160
gaaagtgaag agaaattcgg aagcatggca cattcttctt ttataaata catgcctgac 2220
ttctttttc catgatgatg atatatgcat atgatagata tacaagcaat cttcttcaag 2280

gagtttgaaa ttttgtcttc caggagcaaa aaaaagtttt tttttatata tgtttgtaca 2340
 caagaatagt taccaatitg ctttggcttt acgtgtgca agtttatata gttttcaatt 2400
 tctttgtott tacattttct ttgtcttcta tcttctctca tttagtcttt gggagaatta 2460
 ggaaaaggga gcggaaaggt aagaaatgct tgcgtatttt actaattcgg caaacatcca 2520
 atttggcaaa cagcagcctg tgcaacgctc tcgagatgac agtatctttg attacactct 2580
 aaatctogat gaccogacca aaaagagcga acaagaaat aatcttgtgc attcgaatat 2640
 gatggagat tttttccccc ttattctaaa tgttgacata gcgtgtatgt tataataaaa 2700
 aaaaagaatt gtacaaactt tcttttcttc tctttttatt ttactcttat gtcaatactc 2760
 acttatctgg aatttcactc ctactataca ctacctgtcc ttgcggcatt gtgttggtg 2820
 ctaaagccgt ttcactcaca gcaagacaat ctcaagtata aatttttaat gttgatggcc 2880
 gcctctacgg catcgatttg ggacaattat atcgtttata atcgogcttg gtggtaactgt 2940
 cctacttgby ttgtggctgt cattggctat gtacctctag a 2981

<210> 78

<211> 1/49

<212> DNA

<213> *Blakeslea trispora*

<400> 78

atgtctgac aaaagaagca tattgttgtc attgggtgcc gtattggcgg aactgctact 60
 gctgctogto ttgctogtga aggttttcca gttactgttg ttgaaaagaa cgacttttcc 120
 ggtggccgtt gttcattcat tcatcagat ggtcatcact ttgatcaggg tccctcactc 180
 tattttgatgc ctaagctttt tgaagatgca tttgtgtatt tggatgaacg tattgggtgat 240
 catttggatt tgettcgctg tgacaataac tataaggctc attttgacga cgggtgatgcc 300

gtacaactct ctccgattt aaccaagatg aaggcggaat tggaccgtat tgaggggtccc 360

cttggtattg gtatattctt ggatttcatt aaggaacac atgtccatta tgaacaagggt 420

acatttattg ctatcaagcg caactttgaa accatttggg atttgattcg tcttcagtat 480

gtgcctgaaa tctttcgctt gcacttgttt ggtaagatct atgaccgagc cagtaaatat 540

ttccaaaacca aqaaaatcggt tatggctttt acttttcaaa caatgtatat gggtagtgcg 600

ccttatgatg ctccagcaggt ttacagtttg ttacaatata ccgagtttgc tgaaggatc 660

tggtagcttc gtgtgtgttt caacatggtt gttcagaagc ttgagtctat cgcctccaaa 720

aagtacggtg ctgaattcag atatcaatcg cctgttgcta aaattaacac tgtcgataaa 780

gacaagcgtg taaccgggtt cactttggaa agcggagaag tcattgaagc cgatgcagtc 840

gtatgtaatg cggatcttgt ttatgcttat caccatctgt tacctccttg caattggaca 900

aagaagacat tagcctcaaa gaaactcaat tcacatctta tttcgtttta ttggtccatg 960

tcacaaaagg tgcctcaatt agacgtacac aatatcttct tggctgaagc ctacaaggaa 1020

agttttgatg agattttcaa cgaattcggg ttgcctcttg aagcttcatt ctatgtcaac 1080

gttccatctc gaattgatga atctgccgca cctcccaaca aggactccat tattgtgttg 1140

gttccaattg gccatgatga gagtaagaca ggaacagtg ctgaagaaaa ttatcctgag 1200

ttggtaaacc gtgcacgcaa gatggttctg gaagttagcg aagtcggtt gggagtaaac 1260

aacttttgcta atttgattga acatgaagaa gtgaatgac ctagtgtttg gcaagcaag 1320

tttaaccttt ggagagggtt tattcttggg ctttctcatg atgtgttcca agttctctgg 1380

ttcagaccta gtaccaagga ttccacaaac cgttatgata atctttctt tgtcggagct 1440

agtacacatc caggtagctg tgttccctat gttcttgctg gaagtaagct tacttccgac 1500

caagtctgta aaagctttgg ccagaatccc ttaccaagaa agttacaaga tagccaaaag 1560

aagtatgctc ctgaacaaac tcgtaagacc gaaagccatt ggatctatta ttgtcttgct 1620
tgttactttg ttactttcct ctttttctat ttcttccaa gagatgatac tacaactcct 1680
gcttctttca ttaaccaact tttacctaac gttttccaag tacaagcag caacgatal 1740
cgcatTTaa 1749

<210> 79

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 79

ccgatggcga cgacggaagg ttgtt

25

<210> 80

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 80

catgttcacg cccattgcac cacct

25

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/EP2004/000100A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N1/15 C12N15/80 C12P23/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C12P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98/46772 A (BOVENBERG ROELOF ARY LANS ;GIST BROCADES BV (NL); SELTEN GERARDUS) 22 October 1998 (1998-10-22) the whole document	1-7,9, 14, 19-23, 31-33,35
Y	VAN HEESWIJCK R ET AL: "HIGH FREQUENCY TRANSFORMATION OF MUCOR WITH RECOMBINANT PLASMID DNA" CARLSBERG RESEARCH COMMUNICATIONS, vol. 49, no. 7, 1984, pages 691-702, XP009031431 ISSN: 0105-1938 the whole document	1-3,7, 14, 19-21, 31-34

-/-

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "A" document member of the same patent family

Date of the actual completion of the international search

27 May 2004

Date of making of the international search report

29/06/2004

Name and mailing address of the ISA
European Patent Office, P.B. 6516 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Authorized officer

Vogt, T

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/EP2004/000100

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>NAVARRO EUSEBIO ET AL: "Overexpression of the <i>crgA</i> gene abolishes light requirement for carotenoid biosynthesis in <i>Mucor circinelloides</i>"</p> <p>EUROPEAN JOURNAL OF BIOCHEMISTRY, vol. 267, no. 3, February 2000 (2000-02), pages 800-807, XP002282392</p> <p>ISSN: 0014-2956</p> <p>the whole document</p>	<p>1-3,7, 14, 19-21, 31-34</p>
Y	<p>NAVARRO E ET AL: "A negative regulator of light-inducible carotenogenesis in <i>Mucor circinelloides</i>"</p> <p>MGG MOLECULAR GENETICS AND GENOMICS, vol. 266, no. 3, November 2001 (2001-11), pages 463-470, XP002282393</p> <p>ISSN: 1617-4615</p> <p>the whole document</p>	<p>1-3,7, 14, 19-21, 31-34</p>
Y	<p>SKORY C D: "Homologous recombination and double-strand break repair in the transformation of <i>Rhizopus oryzae</i>."</p> <p>MGG MOLECULAR GENETICS AND GENOMICS, vol. 268, no. 3, November 2002 (2002-11), pages 397-406, XP002282394</p> <p>ISSN: 1617-4615 (ISSN print)</p> <p>the whole document</p>	<p>1-3,7, 14, 19-21, 31-34</p>
A	<p>US 6 413 736 B1 (JOLLY SETSUO OMATA ET AL) 2 July 2002 (2002-07-02)</p> <p>the whole document</p>	
A	<p>US 2002/051998 A1 (SCHMIDT-DANNERT CLAUDIA ET AL) 2 May 2002 (2002-05-02)</p> <p>the whole document</p>	
A	<p>LEE P C ET AL: "Metabolic engineering towards biotechnological production of carotenoids in microorganisms."</p> <p>APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, vol. 60, no. 1-2, October 2002 (2002-10), pages 1-11, XP002282395</p> <p>ISSN: 0175-7598</p>	
A	<p>ZHIQIANG A ET AL: "New cosmid vectors for library construction, chromosome walking and restriction mapping in filamentous fungi"</p> <p>GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES, ELSEVIER SCIENCE PUBLISHERS, BARKING, GB, vol. 176, no. 1-2, 17 October 1996 (1996-10-17), pages 93-96, XP004070214</p> <p>ISSN: 0378-1119</p>	

-/-

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/EP2004/000100

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, Y	NICOLAS FRANCISCO E ET AL: "Two classes of small antisense RNAs in fungal RNA silencing triggered by non-integrative transgenes." EMBO (EUROPEAN MOLECULAR BIOLOGY ORGANIZATION) JOURNAL, vol. 22, no. 15, 1 August 2003 (2003-08-01), pages 3983-3991, XP002282396 ISSN: 0261-4189 (ISSN print) the whole document _____	1-3, 31

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PC1/EP2004/000100

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9846772	A	22-10-1998	AU 7642298 A	11-11-1998
			BR 9808859 A	01-08-2000
			CN 1257546 T	21-06-2000
			WO 9846772 A2	22-10-1998
			EP 0979294 A2	16-02-2000
			JP 2001518798 T	16-10-2001
			PL 336345 A1	19-06-2000
			US 6432672 B1	13-08-2002
US 6413736	B1	02-07-2002	US 6015684 A	18-01-2000
			US 5922560 A	13-07-1999
			US 5466599 A	14-11-1995
			US 2003049241 A1	13-03-2003
			AU 688280 B2	12-03-1998
			AU 5008693 A	08-11-1994
			EP 0708604 A1	01-05-1996
			JP 8508885 T	24-09-1996
			NO 954052 A	11-12-1995
			NZ 255742 A	26-05-1997
			WO 9423594 A1	27-10-1994
US 2002051998	A1	02-05-2002	AU 2257401 A	18-06-2001
			EP 1238068 A1	11-09-2002
			WO 0142455 A1	14-06-2001

INTERNATIONALER RECHERCHENBERICHT

Internationaler Aktenzeichen
PCT/EP2004/000100

A. KLASSIFIZIERUNG DES ANMELDUNGSBEFUGNISSTANDES
IPK 7 C12N1/15 C12N15/80 C12P23/00

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierte Mindestprüfgröße (Klassifikationssystem und Klassifikationssymbole)
IPK 7 C12N C12P

Recherchierte aber nicht zum Mindestprüfgröße gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der letzten fünf Jahre Rechercheur konsultierte elektronische Datenbanken (Namen der Datenbanken und evtl. verwendete Suchbegriffe)

EPO-Internal, BIOSIS

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Detr. Anspruch Nr.
X	WO 98/46772 A (BOVENBERG ROELOF ARY LANS ;GIST BROCADES BV (NL); SELTEN GERARDUS) 22. Oktober 1998 (1998-10-22) das ganze Dokument	1-7,9, 14, 19-23, 31-33,35
Y	VAN HEESWIJCK R ET AL: "HIGH FREQUENCY TRANSFORMATION OF MUCOR WITH RECOMBINANT PLASMID DNA" CARLSBERG RESEARCH COMMUNICATIONS, Bd. 49, Nr. 7, 1984, Seiten 691-702, XP009031431 ISSN: 0105-1938 das ganze Dokument	1-3,7, 14, 19-21, 31-34

-/-

☒ Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

☒ Siehe Anhang Patentfamilie

* Bezeichnen Kategorien von angegebenen Veröffentlichungen :

"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besondere Bedenken anzusehen ist
"B" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie auszuführen)

"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausbeutung oder andere Maßnahmen bezieht

"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem besprochenen Prioritätsdatum veröffentlicht worden ist

"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

"X" Veröffentlichung von besonderer Bedeutung: die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf andererlicher Fähigkeit beruhend betrachtet werden

"Y" Veröffentlichung von besonderer Bedeutung: die beanspruchte Erfindung kann nicht als auf erfindungsfähiger Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

"Z" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

27. Mai 2004

Abgeschlossenheit des internationalen Recherchenberichts

29/06/2004

Name und Postanschrift der Internationalen Recherchenbehörde
Europäisches Patentamt, P.O. 5516 Palenstein 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 sp nl,
Fax (+31-70) 340-3016

Bewilligter Dienstleister

Vogt, T

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP2004/000100

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
Y	<p>NAVARRO EUSEBIO ET AL: "Overexpression of the crga gene abolishes light requirement for carotenoid biosynthesis in Mucor circinelloides" EUROPEAN JOURNAL OF BIOCHEMISTRY, Bd. 267, Nr. 3, Februar 2000 (2000-02), Seiten 800-807, XP002282392 ISSN: 0014-2956 das ganze Dokument</p>	<p>1-3,7, 14, 19-21, 31-34</p>
Y	<p>NAVARRO E ET AL: "A negative regulator of light-inducible carotenogenesis in Mucor circinelloides" MGG MOLECULAR GENETICS AND GENOMICS, Bd. 266, Nr. 3, November 2001 (2001-11), Seiten 463-470, XP002282393 ISSN: 1617-4615 das ganze Dokument</p>	<p>1-3,7, 14, 19-21, 31-34</p>
Y	<p>SKORY C D: "Homologous recombination and double-strand break repair in the transformation of Rhizopus oryzae." MGG MOLECULAR GENETICS AND GENOMICS, Bd. 268, Nr. 3, November 2002 (2002-11), Seiten 397-406, XP002282394 ISSN: 1617-4615 (ISSN print) das ganze Dokument</p>	<p>1-3,7, 14, 19-21, 31-34</p>
A	<p>US 6 413 736 B1 (JOLLY SETSUKO OMATA ET AL) 2. Juli 2002 (2002-07-02) das ganze Dokument</p>	
A	<p>US 2002/051998 A1 (SCHMIDT-DANNERT CLAUDIA ET AL) 2. Mai 2002 (2002-05-02) das ganze Dokument</p>	
A	<p>LEE P C ET AL: "Metabolic engineering towards biotechnological production of carotenoids in microorganisms." APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, Bd. 60, Nr. 1-2, Oktober 2002 (2002-10), Seiten 1-11, XP002282395 ISSN: 0175-7598</p>	
A	<p>ZHIQIANG A ET AL: "New cosmid vectors for library construction, chromosome walking and restriction mapping in filamentous fungi" GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES, ELSEVIER SCIENCE PUBLISHERS, BARKING, GB, Bd. 176, Nr. 1-2, 17. Oktober 1996 (1996-10-17), Seiten 93-96, XP004070214 ISSN: 0378-1119</p>	

-/-

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen
PCT/EP2004/000100

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
P,Y	<p>NICOLAS FRANCISCO E ET AL: "Two classes of small antisense RNAs in fungal RNA silencing triggered by non-integrative transgenes."</p> <p>EMBO (EUROPEAN MOLECULAR BIOLOGY ORGANIZATION) JOURNAL, Bd. 22, Nr. 15, 1. August 2003 (2003-08-01), Seiten 3983-3991, XP002282396 ISSN: 0261-4189 (ISSN print) das ganze Dokument</p>	1-3,31

INTERNATIONALER RESEARCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Abkürzungen

PCT/EP2004/000100

Im Recherchenbericht angeführtes Patentedokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 9846772 A	22-10-1998	AU 7642298 A	11-11-1998
		BR 9808859 A	01-08-2000
		CN 1257546 T	21-06-2000
		WO 9846772 A2	22-10-1998
		EP 0979294 A2	16-02-2000
		JP 2001518798 T	16-10-2001
		PL 336345 A1	19-06-2000
		US 6432672 B1	13-08-2002
US 6413736 B1	02-07-2002	US 6015684 A	18-01-2000
		US 5922560 A	13-07-1999
		US 5466599 A	14-11-1995
		US 2003049241 A1	13-03-2003
		AU 688280 B2	12-03-1998
		AU 5008693 A	08-11-1994
		EP 0708604 A1	01-05-1996
		JP 8508885 T	24-09-1996
		NO 954052 A	11-12-1995
		NZ 255742 A	26-05-1997
US 2002051998 A1	02-05-2002	WO 9423594 A1	27-10-1994
		AU 2257401 A	18-06-2001
		EP 1238068 A1	11-09-2002
		WO 0142455 A1	14-06-2001